Borough Data Partnership

Meeting #4

Data Science Institute
Imperial College London

7 October 2015

Agenda

13.30 **Tea/coffee reception**

14.00 **Welcome & Introduction**
Andrew Collinge, Assistant Director of GLA Intelligence Unit

GLA news and projects update, Andrew Collinge
H2020; LDS strategy; ODI award…and more

Cabinet Office, Joshan Meenowa
Data Science in Government

University College London (CASA), Pete Ferguson
Project Sync: City modelling

15.00 break/refreshments

Agenda

15.10 – 15.55 Data viz/sharing sessions:

Data Observatory	Data Sharing
Data Science Institute – demo of the studio	TfL, Freight data
UCL - City modelling: Project Sync	London Councils, cross-border data harmonisation
GLA - web maps and data models	GLA, pan-London data projects

15.55 – 16.40 Repeat the above

16.45 Next steps & close

Data Sharing

1. **New York Trash**

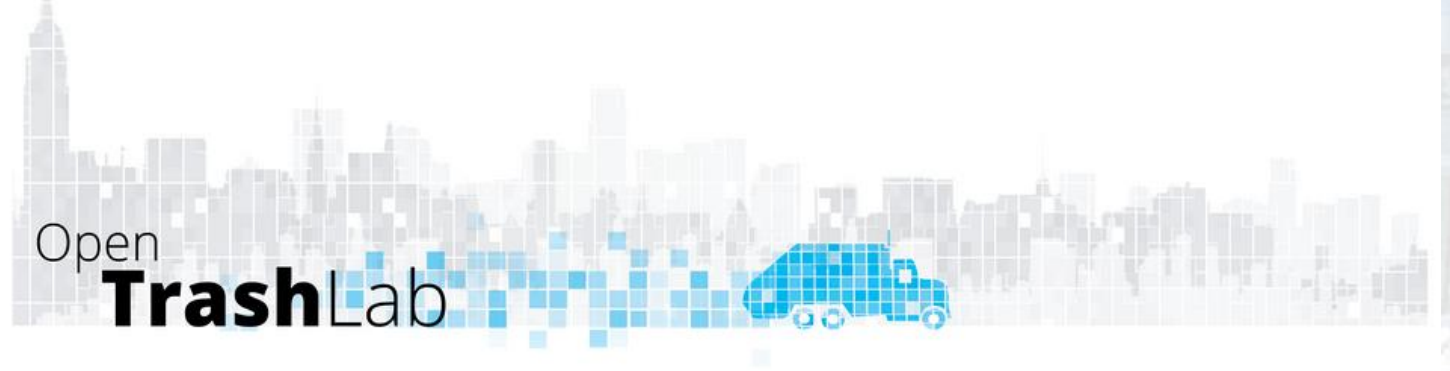
2. **Pan-London Freight data**
   Simon Pitman, Project Manager Freight and Fleet, TfL

3. **Cross-border data Harmonisation**
   Daniel Quirk, Principal Policy Officer, London Councils

4. **Pan-London datasets**
   Alan Lewis, Senior R&S Analyst, GLA

Data Sharing

Example: Urban Waste Analytics: NYC/CUSP

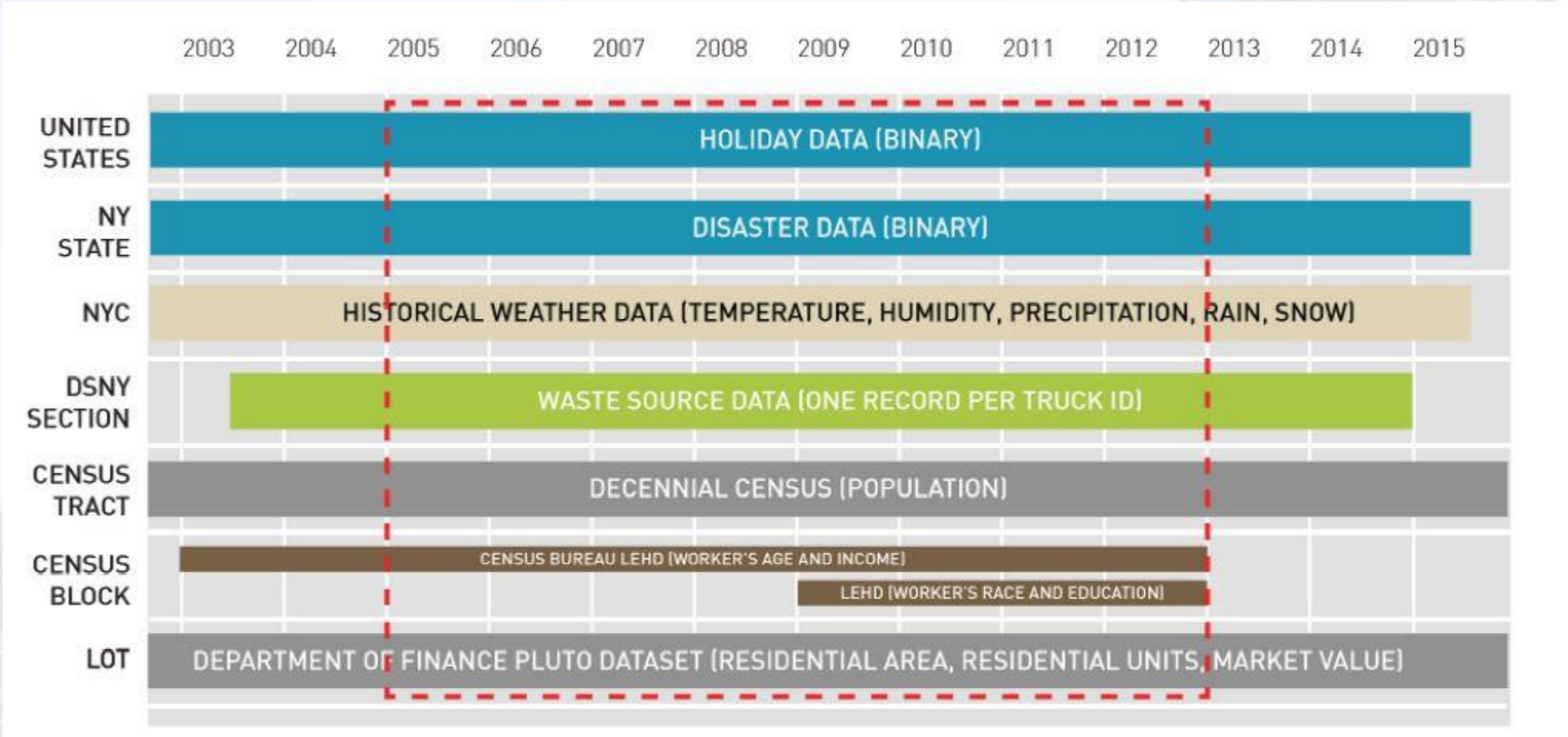


Data Sharing

- NYC municipal waste amounts to 11 million pounds of waste daily (17% recycled)
- Aims: help plan logistics of waste collection (short-term) and longer-term planning
- Objective: predict weekly residential waste generation for DSNY sections

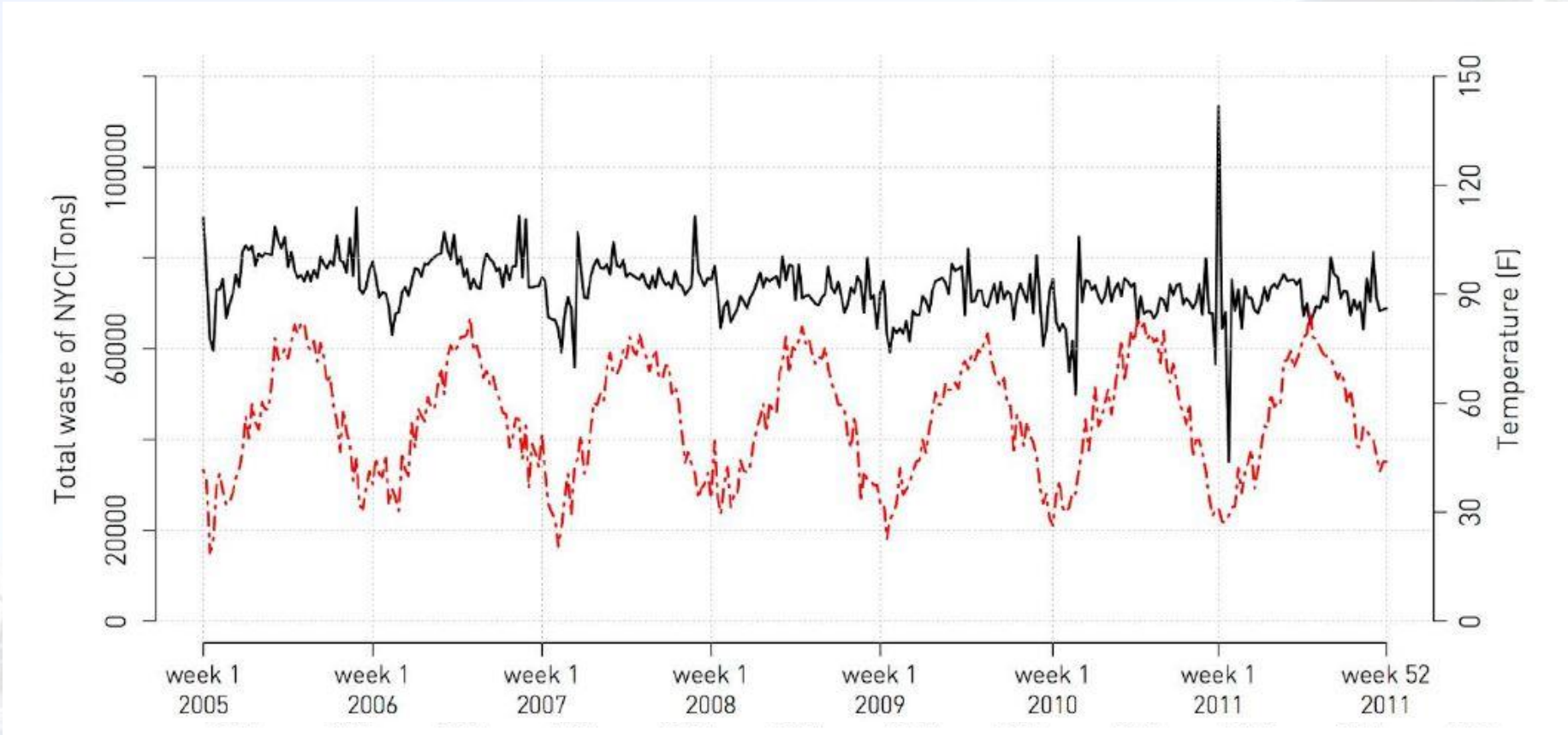
Data Sharing

- the data



Data Sharing

- **trends**



Data Sharing

- **Maths and machine learning**

$$\epsilon = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log(p_i + 1) - \log(a_i + 1))^2}$$

- Refuse:
  $R^2 = 0.97$, RMSLE = 0.05
- Paper, magazine and cardboard:
  $R^2 = 0.8$, RMSLE = 0.18
- Plastic, metal and glass:
  $R^2 = 0.9$, RMSLE = 0.089

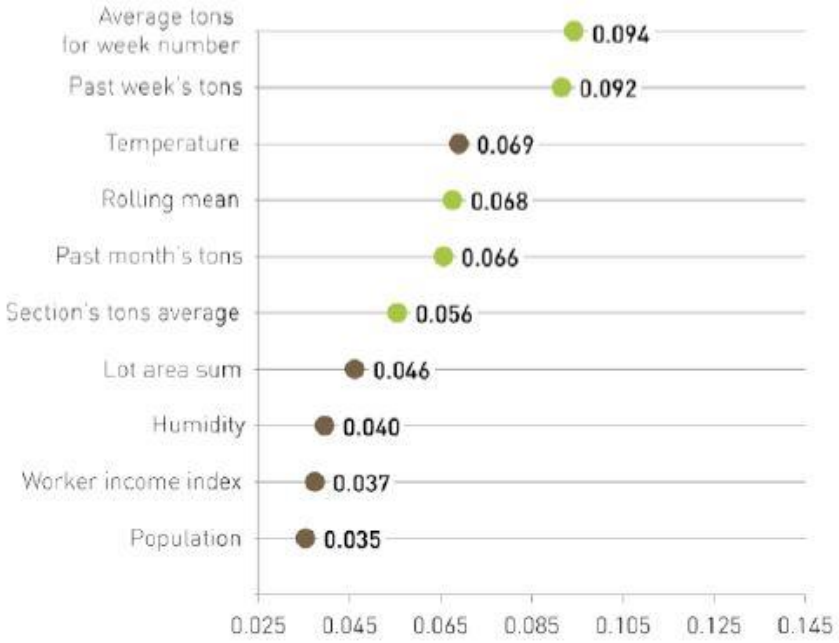
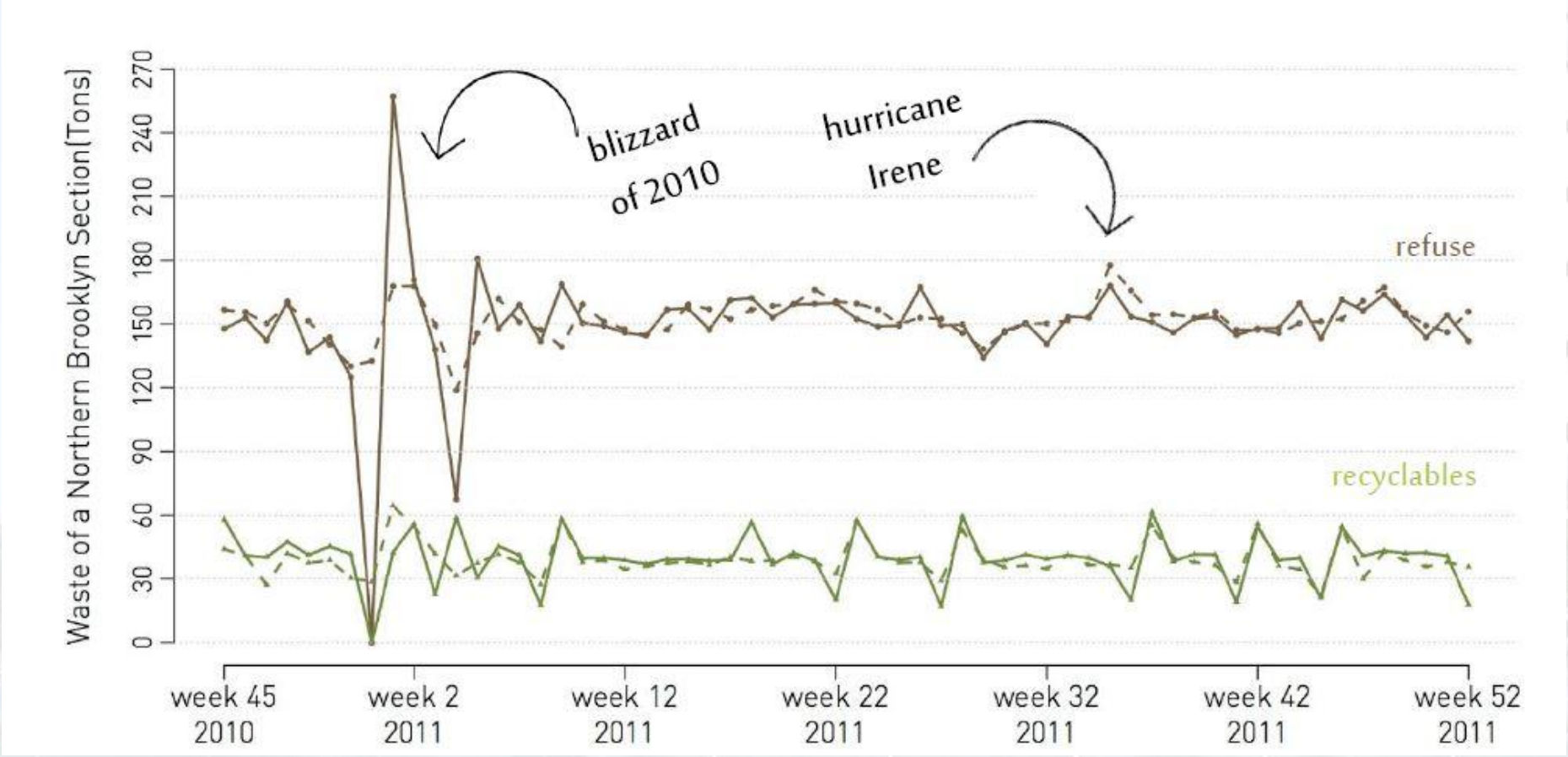


Fig. 1 - Feature relative importance for **refuse**

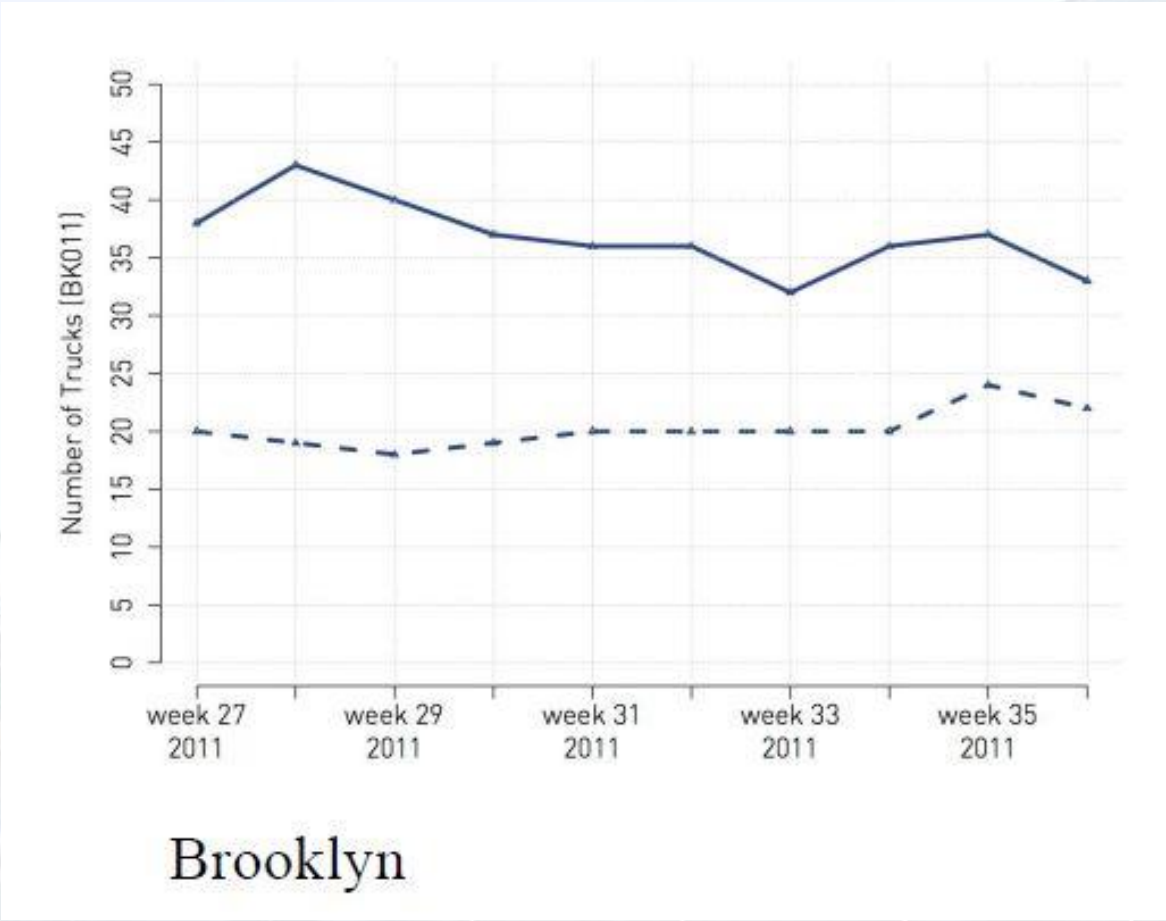
Data Sharing

- **Predictions**



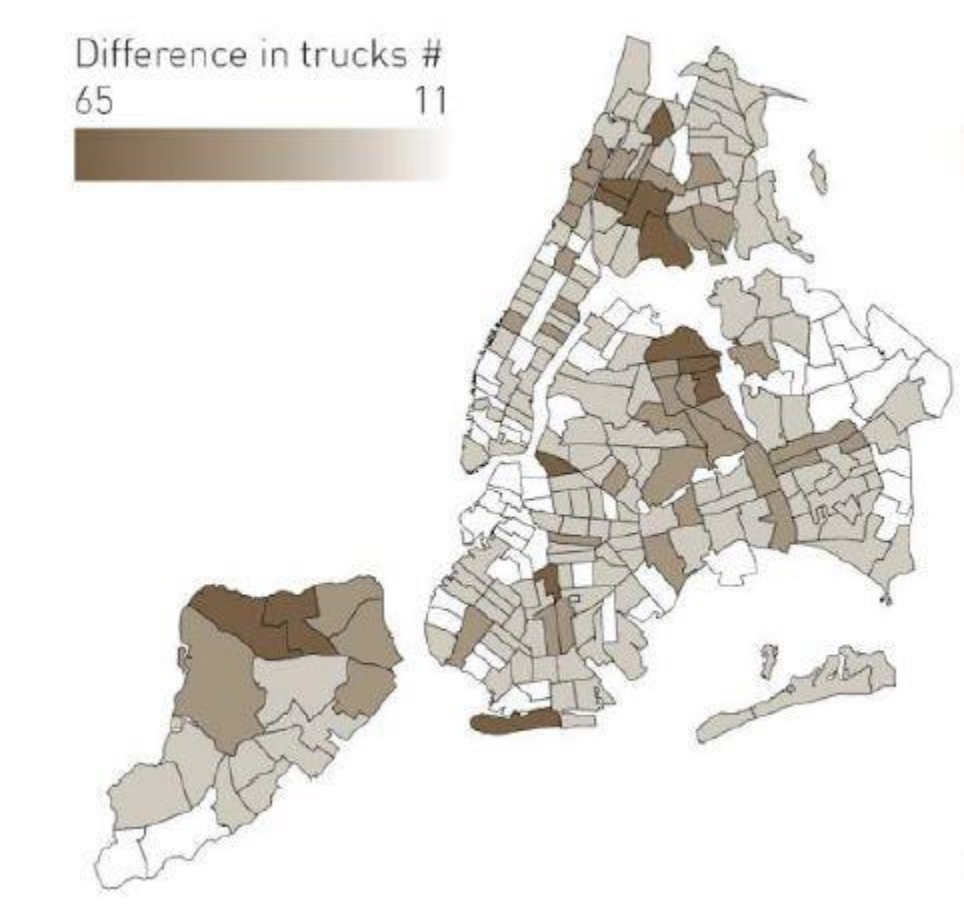
Data Sharing

- **Predictions**

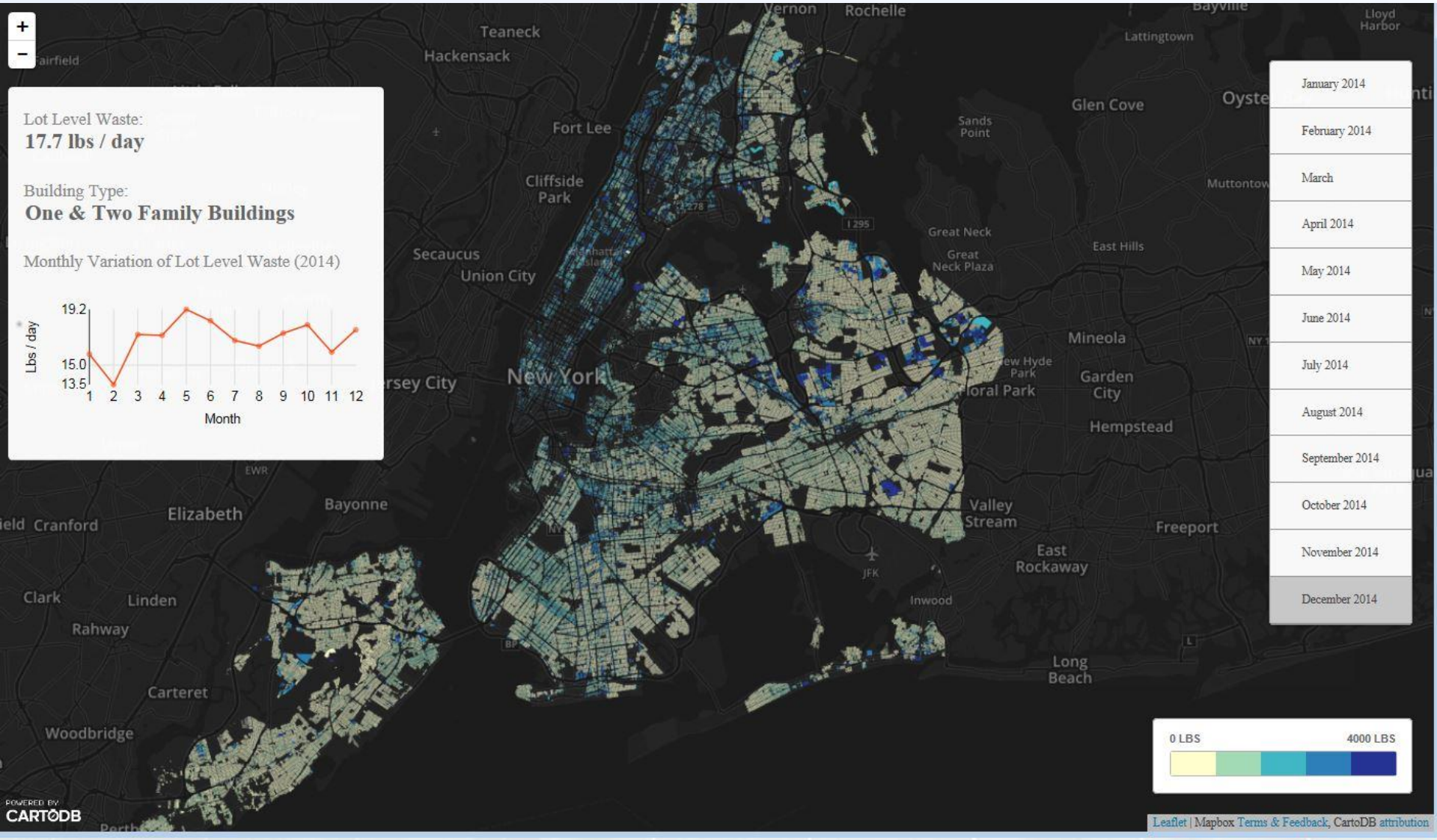


Data Sharing

- **Too many trucks?**



Data Sharing



http://opentrashlab.cloudapp.net/Lots_all_tool/lots_2014.html

Data Sharing

- thoughts/reactions?
- of interest or not?
- do we have the data?
- could we get the data?
- could we collect the data in future?

Data Sharing

1. New York Trash

2. **Pan-London Freight data**
   Simon Pitman, Project Manager Freight and Fleet, TfL

3. **Cross-border data Harmonisation**
   Daniel Quirk, Principal Policy Officer, London Councils

4. **Pan-London datasets**
   Alan Lewis, Senior R&S Analyst, GLA

Data Sharing

Prioritised pan-London datasets from last time:

Themes	Top Pan-London datasets:
Transport	Bike storage/parking
	Traffic by street/Traffic count data
Environment	Carbon offset
Finance	Cost of consultancy fees
Education and Youth	School capacity (waiting)/roll data (Pupil flows from home to school)
	Missing Children across London and associated trends
Buildings and planning	High rise/Sheltered housing - property use (Building)
	Empty buildings/sites - council tax/revenue information/planning and development
	Shop Frontage Data - % vacancies, occupiers, non-retail
Wellbeing	Household income estimates
Health	GP's/Health service - Performance, Social care availability - Linking ASC & data
	Health/GP data (Trusted connection)

Data Sharing

Harvesting data from LB datastores:

Camden

BARNET
LONDON BOROUGH

London Borough of Redbridge

London Borough of Hounslow

Others?

www.londoncouncils.gov.uk

LONDON COUNCILS

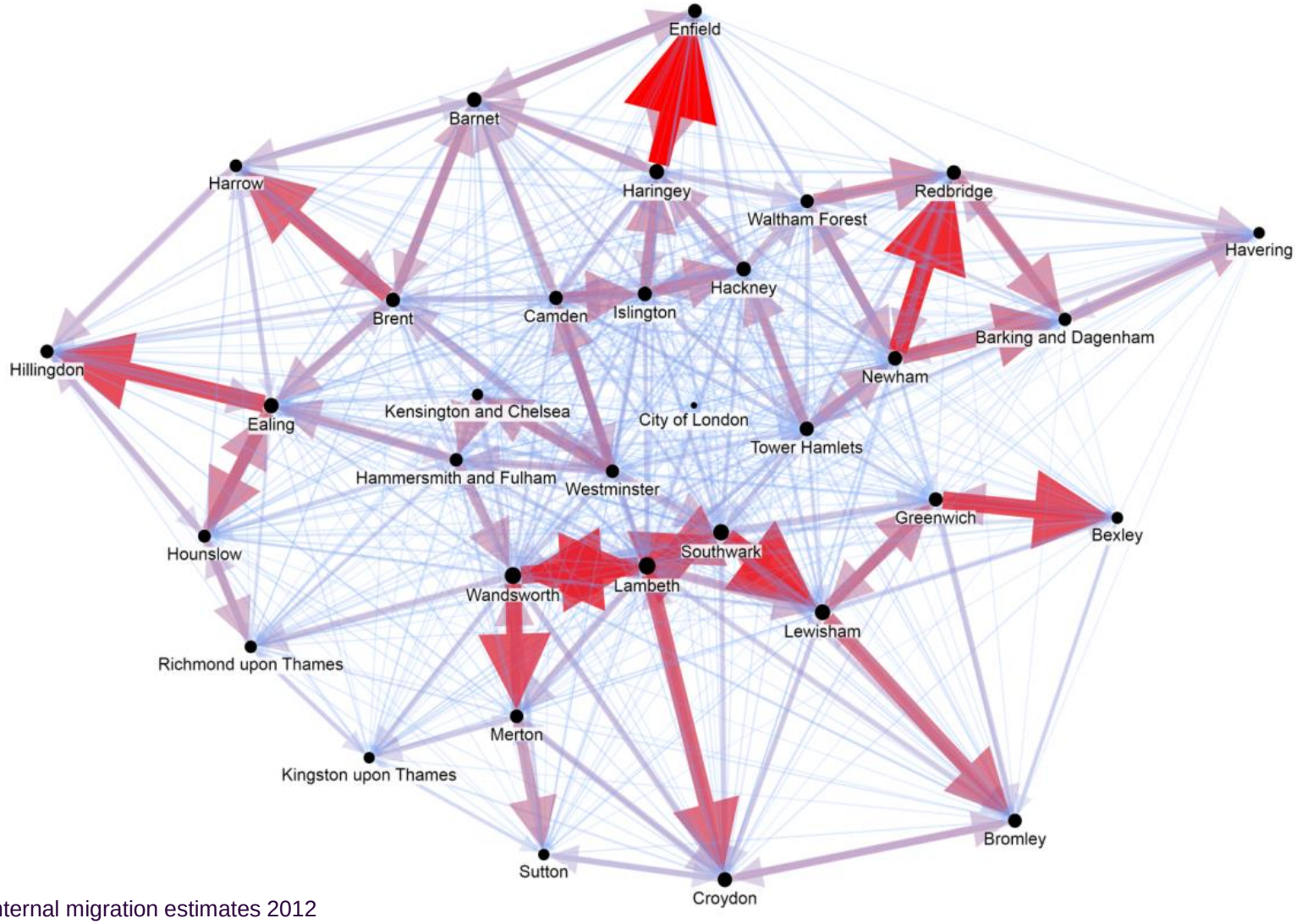
CROSS BORDER MOBILITY

a live example of data harmonisation

October 2015

A question about mobility

- Executive driven, pan-London work
- Understand the flow of people across local authority borders
- Focus on ‘vulnerable’ groups who present a significant cost to public services
- Understand the broader trends that are shaping change and churn across the capital



ONS internal migration estimates 2012

Cross-border solutions for cross-border themes

- Looking outside of our authority borders to better understand cross-border issues such as mobility, welfare reform, homelessness
- Some of the challenges we face are in consequence of a much wider ecology beyond our borough boundaries: London
- Sometimes we need to build intelligence that reflects the region, to better understand our own role within pan-London space

A problem with no answer

- However data to explore these cross-border themes is not readily available or accessible at a pan-London level
- Sensitive data brings challenges of its own; both political and legal
- Mixed data landscape across London authorities
- Boroughs facing the most challenging financial settlement yet- in-house expertise and capacity stretched and in some cases depleting
- By no means a simple task

How do we bring data together?

- We don't always realise the full potential of our data
  - infrastructure, strategy, capacity, legacy
- Unlocking incredibly rich data sources within our reach
- GLA Intelligence offers a dedicated intelligence function positioned for city/big data ambitions
- GLA Intelligence/ London Councils as an extension of borough capability and capacity

Opportunity to pilot a data harmonisation task with the support of GLA Intelligence, bringing relevant datasets together from three boroughs (initially) to:

a) better understand cross-border movement

b) unearth issues around the act of data harmonising as well as further insights that can be driven from the exercise

Produce a closed-loop, securely harmonised piece of intelligence for internal use- facilitated by expertise within the London local government family

Fact finding and data audit

- Point of contact appointed in each borough acting as the information gatherer
- Brought together in a series of workshops to explore potential datasets ripe for harmonisation

CHILDREN, YOUNG PEOPLE & FAMILIES
- CHILDREN UNDER A PROTECTION PLAN
- CHILD IN NEED PLANS
- LOOKED AFTER CHILDREN (PLACED)
- YOUTH OFFENDING
- WORKING WITH FAMILIES (TROUBLED FAMILIES)

ADULT SOCIAL CARE AND HEALTH
- SUPPORTED LIVING AND CARE ACCOMMODATION
- ADULTS WITH COMPLEX SUPPORT NEEDS

HOUSING
- HOUSING BENEFIT AND WELFARE ASSISTANCE
- STATUTORY HOMELESSNESS
- TEMPORARY & EMERGENCY ACCOMMODATION
- HOUSED 16-17 YEAR OLDS
- SOCIAL HOUSING SECTOR
- PRIVATE HOUSING SECTOR
- SUPPORTED ACCOMMODATION
- ROUGHSLEEPING

COMMUNITY SAFETY
- GANGS AND COUNTY LINES
- YOUTH JUSTICE
- DOMESTIC AND SEXUAL VIOLENCE

CUSTOMER SERVICE CENTRE AND COUNCIL TAX

Critical challenges and early lessons

- Data quality
- Interoperability of datasets between boroughs
- Unique identifying data e.g. NINO
- In-house intelligence coordination and resource
- Central government standards and ontologies
- Perceptions of data protection and the legislative framework
- Rhetoric- perception of data sharing is not accurate for harmonisation

- Quickly apparent not so ripe
- Unique identifier probably singlehandedly the most limiting factor
- Narrowed down to most viable options. This narrowed list now considered a range of factors…

Type of data	Executive interest	Officer interest	Unique unifying identifier	Ease of extraction	Data quality	Sensitivity	Data sharing restrictions
Housing Benefit	HIGH	HIGH	YES	GOOD	GOOD	HIGH	SSAA significant
Placement of children looked after by LA	HIGH	MEDIUM-LOW	NO	HARD	MIXED	SUPER HIGH	Challenges
School admissions	MEDIUM	MEDIUM	YES	GOOD if intercepting returns	GOOD	HIGH	Not without assurances
FE/HE learners	MEDIUM	MIXED	YES- for recent cohorts	HARD	POOR	FAIR	Could be problematic

- Housing benefit presents us with an opportunity to use NINOs to study the spatial trends of claims over time; tenure type can also be included in this
- A sense we must hold our nerve with housing benefit data; this is by far the most valuable of datasets to harmonise
- It is also the most viable – HOWEVER, currently restricted by the SSAA and DWP guidance
- Mixed response from the HB teams in the three boroughs in terms of sharing SHBE
- Awaiting guidance from information governance colleagues
- Taking proposal to housing benefit managers group

- Whatever the immediate outcome, data governance and clearly defined processes must underpin any harmonisation arrangement to instil confidence and offer assurance.

- If the immediate sense is there is no legal gateway to support this arrangement, then we will look to work with HB managers to identify current obstacles and set out a practical way forward, assuming wide support from boroughs

Opportunities

- Strategic planning and policy implications
- Making a case for London and the unique challenges it faces
- Articulating the increasing pressures and stresses that demographic change is having on London and its communities
- Fuller, joined-up evidence base
- Longitudinal overview to increase our understanding of pan-London trends
- Inter borough dependencies
- Symbolic: working towards a collective London solution together

Big challenges ahead

- Central government and departments
- Legislative framework
- Getting to grips with information governance
- Joined up, strategic approaches
- Digital platforms
- Investment
- Vision

Making Digital Government Work for Everyone- Digital Government Review

‘It notes that the legal landscape is highly complicated, but also the view of the Information Commissioner’s Office that problems with data sharing are generally cultural and based on misunderstandings of the law, inter-organisational distrust, budgetary restraints, incompatible ICT systems and so on.’

Introducing
SYNC

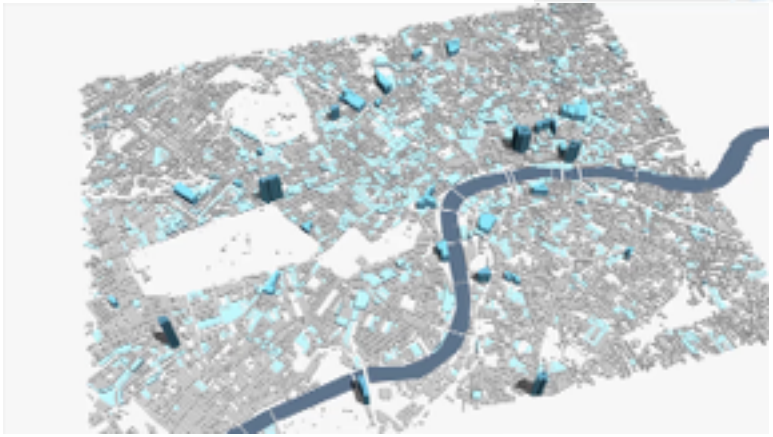
An online collaborative decision support platform for urban planning, design and management

London Borough Data Partnership

7th October 2015
Imperial College London

Who are we and where have we come from?

The Bartlett Centre for Advanced Spatial Analysis

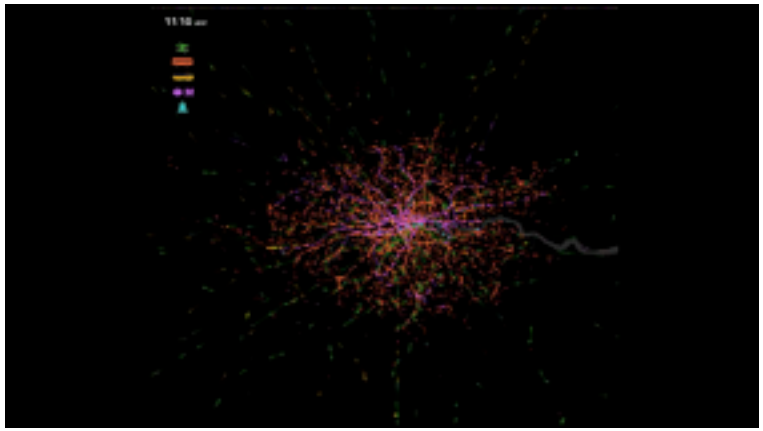


Urban Research Lab at University College London

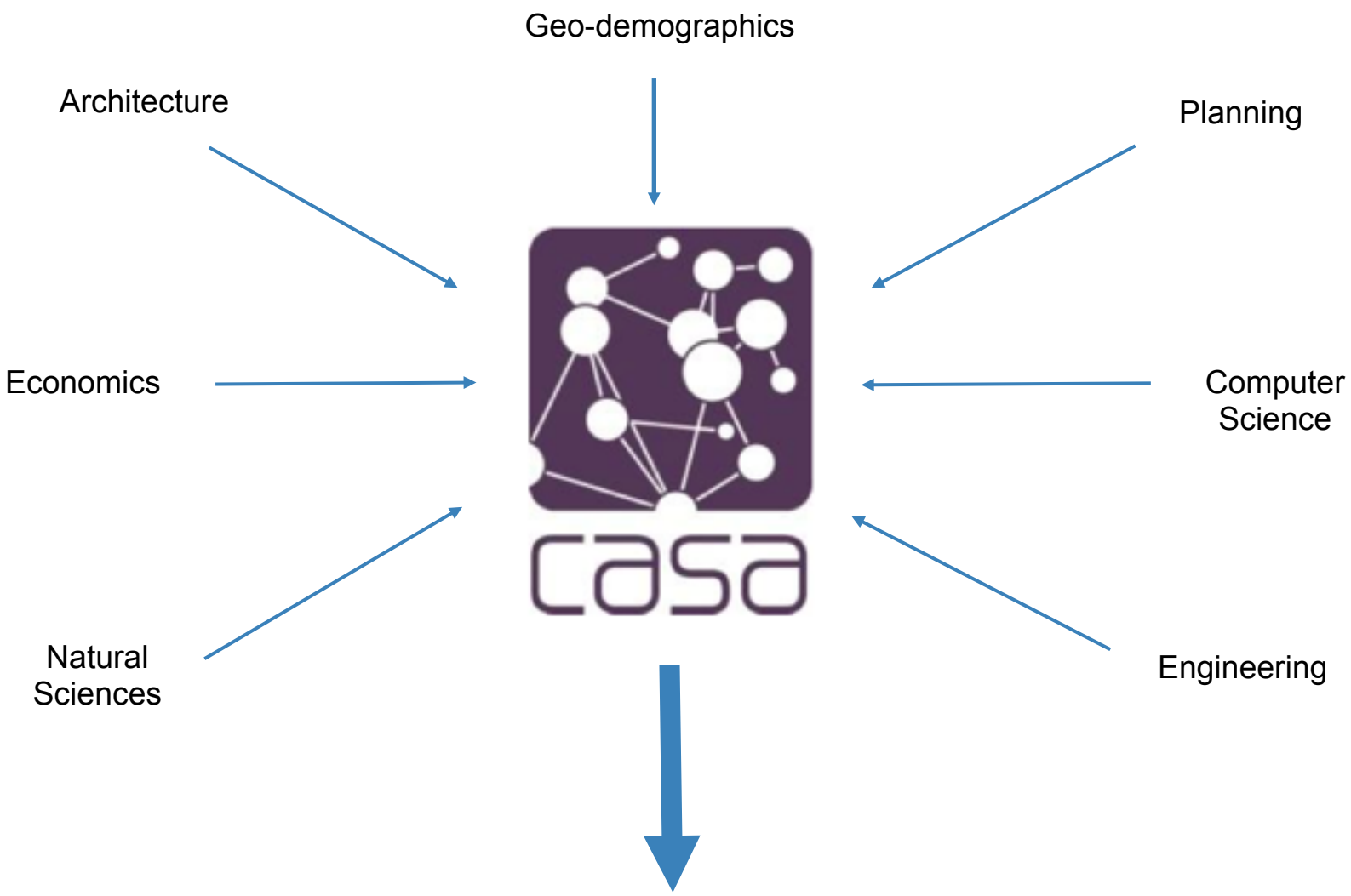
23 PhD Students
24 Research Associates
6 Lecturers
3 Professors

35 Masters Students

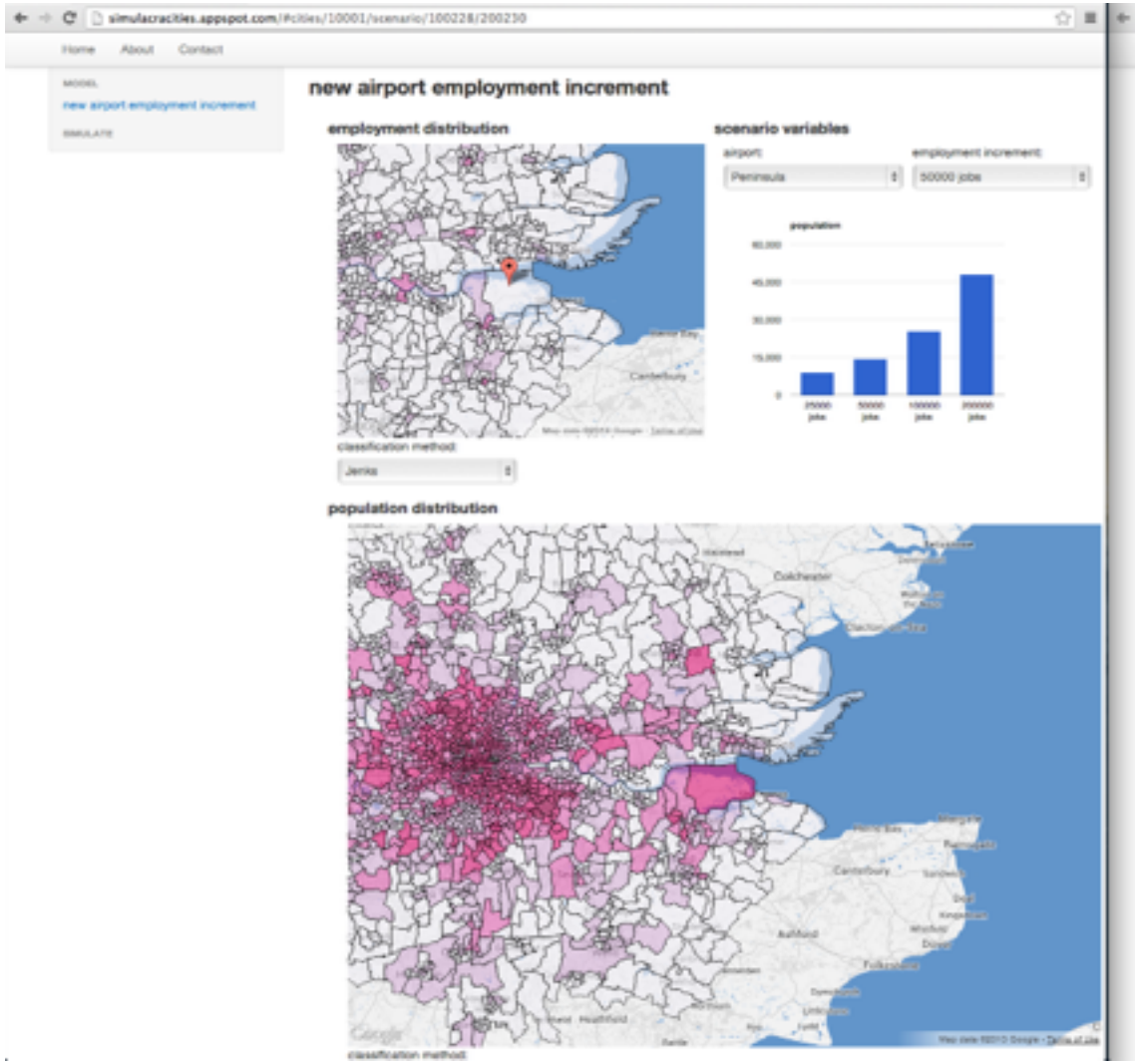
MSc Smart Cities and Urban Analytics
MRes Advanced Spatial Analysis and Visualisation

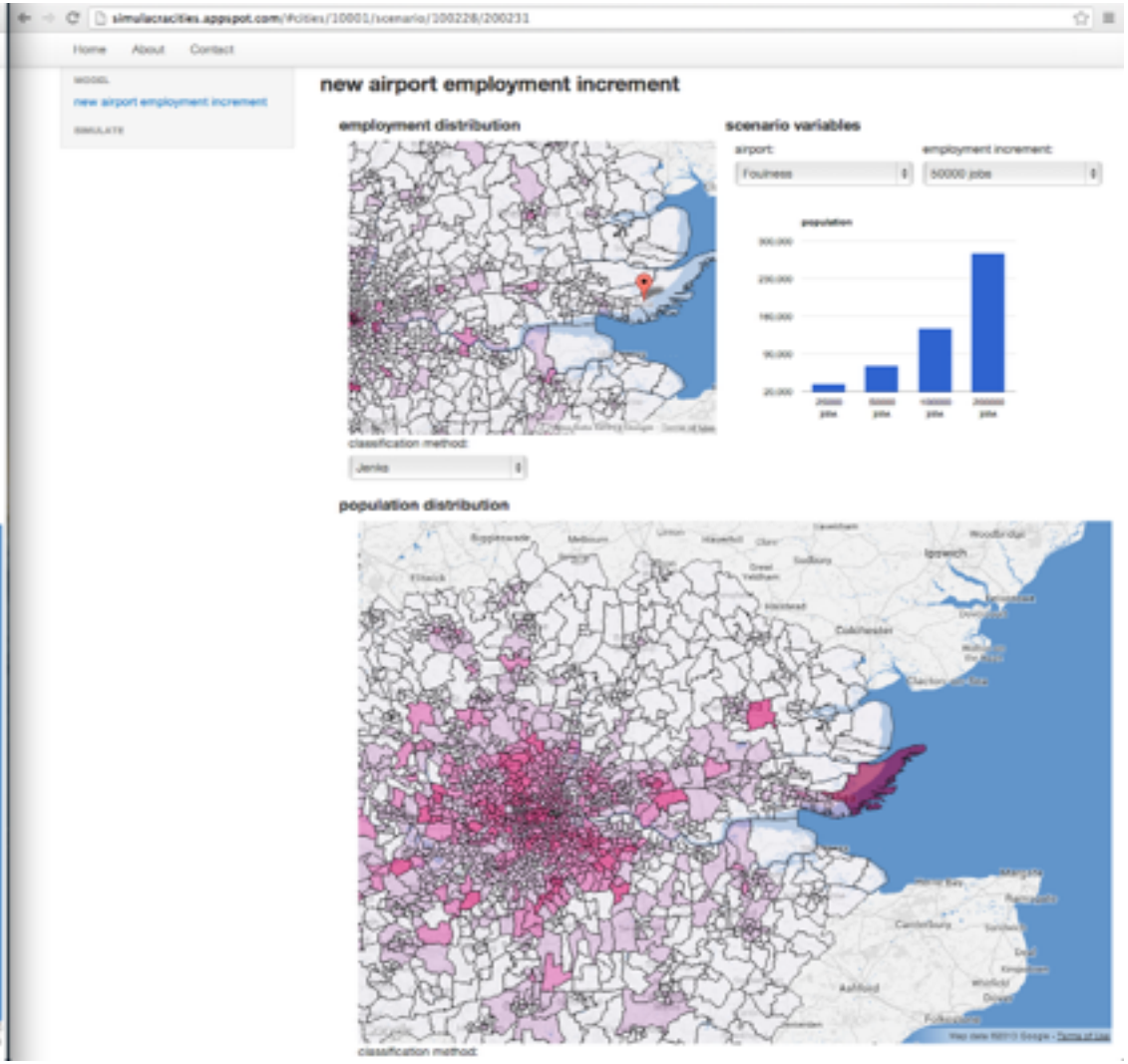


Who are we and where have we come from?



Who are we and where have we come from?





Who are we and where have we come from?

SYNC

Innovate UK

The Challenge

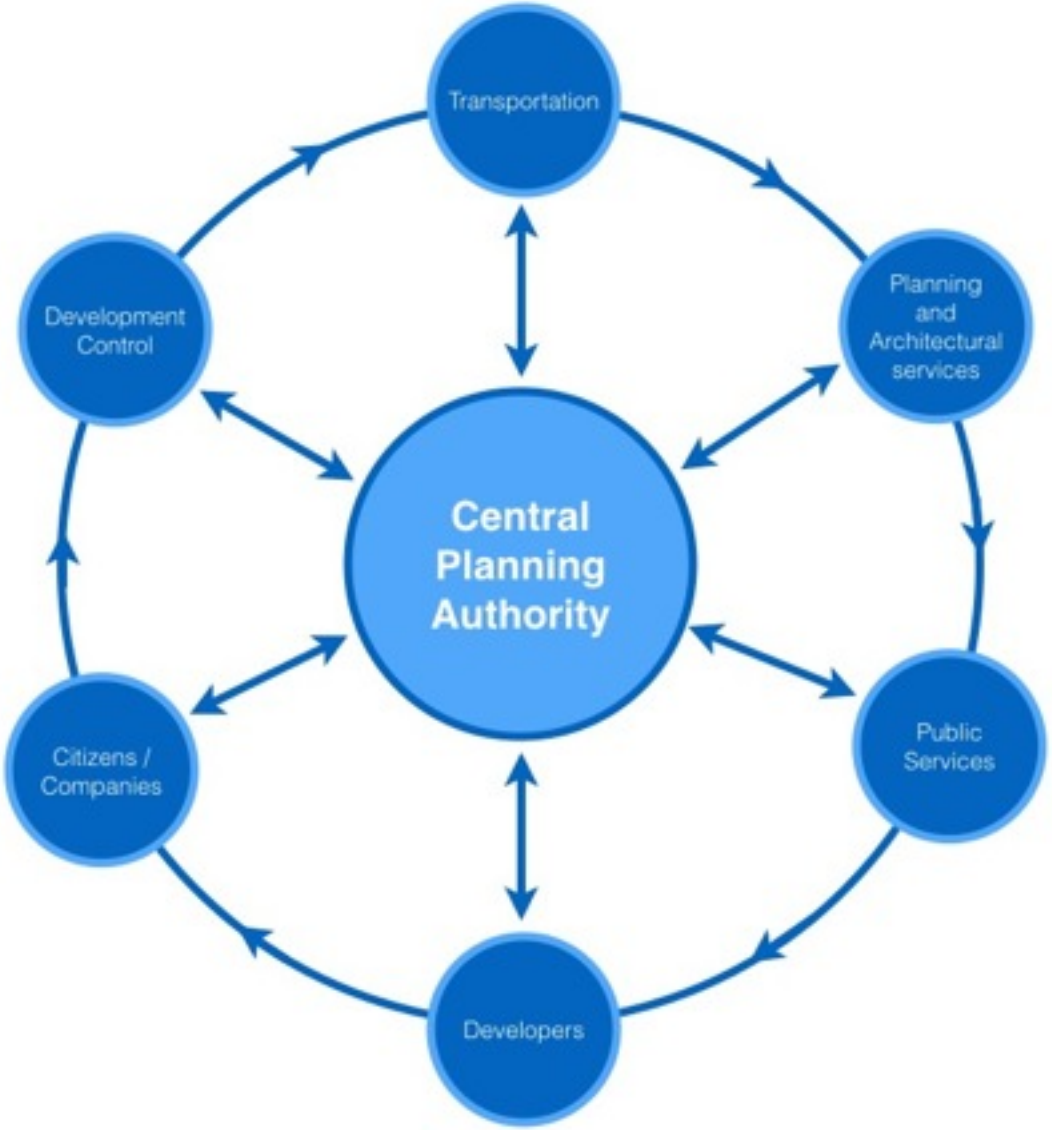
Utilise cloud resources to enable a much wider range of people to benefit from advanced urban modelling applications

Speed up, simplify and automate model processes so that decision makers can operate applications directly using interactive web tools

Enable stakeholders to test, explore and share their plans for the future of the city

The Challenge

We want to support more iterative design and plan formation and consequently more continuous coordination between key city stakeholders



SYNC

Our aim is to enable faster, more contextual evaluation of individual planning policies and design proposals

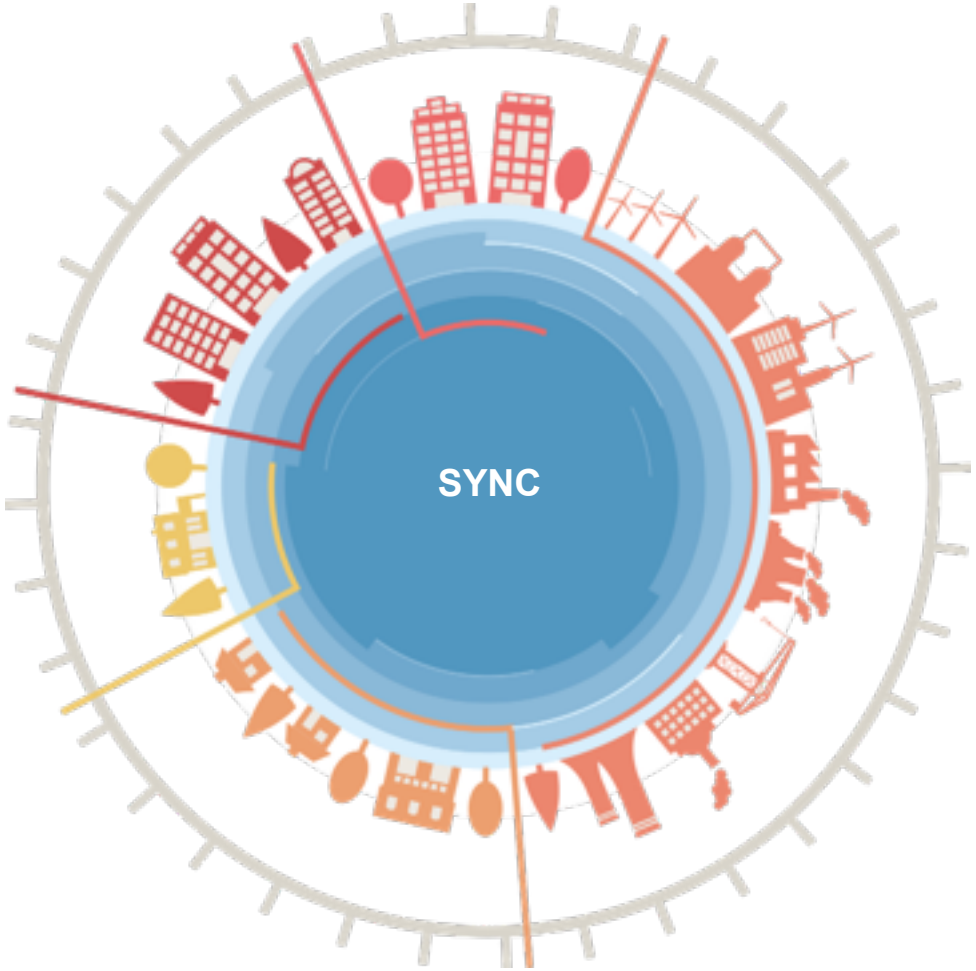
SYNC will open up access to powerful urban modelling applications that encapsulate the interdependency between urban systems

SYNC will provide planning and design tools that facilitate integrated and collaborative decision making between key city stakeholders.

SYNC will enable domain specialists to scale up and connect their applications.

What is SYNC?

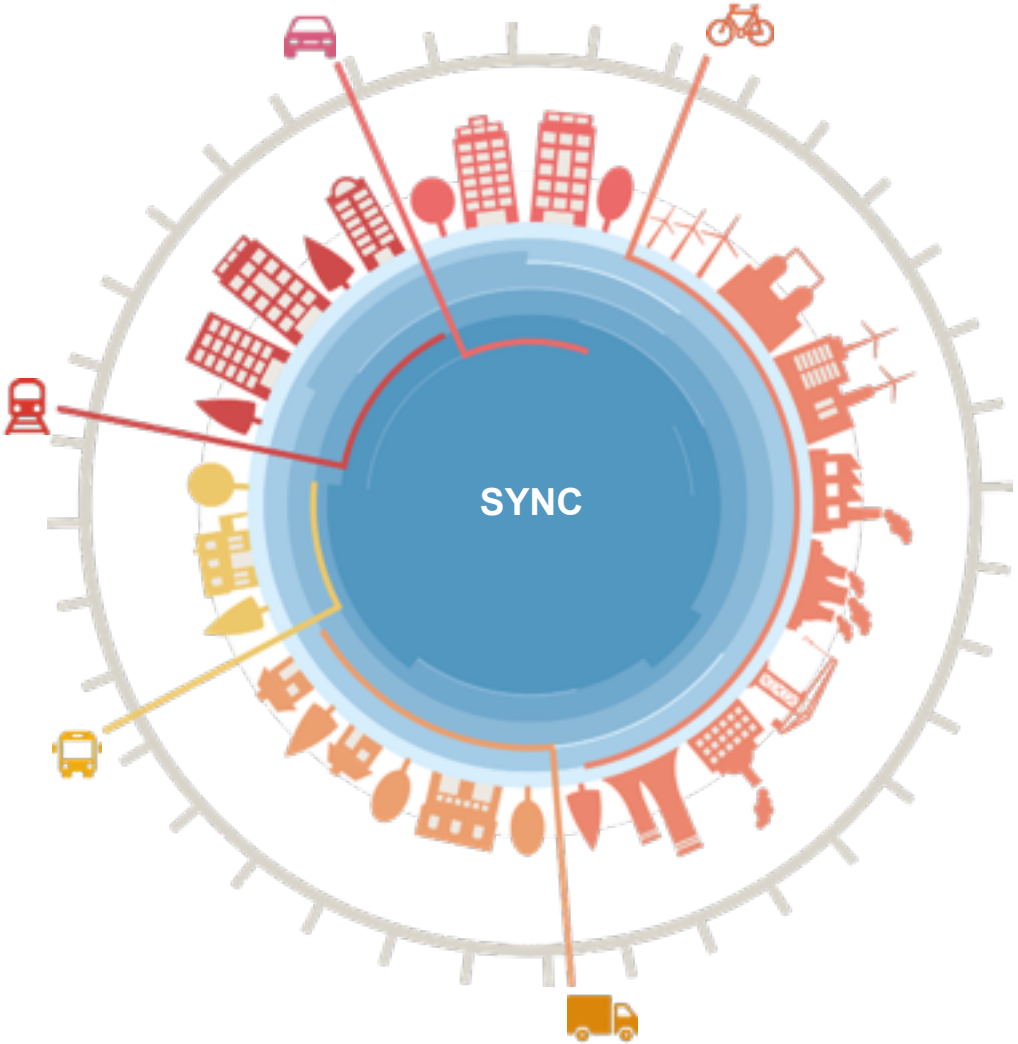
What is SYNC?



SYNC internalises the interdependencies between systems embedded in space. It does this by calculating the benefits and costs of interaction between people, places, infrastructure and services.

SYNC disaggregates these interactions across space so that the impacts of a change to any one element of the urban system in any one location can be evaluated across the city region

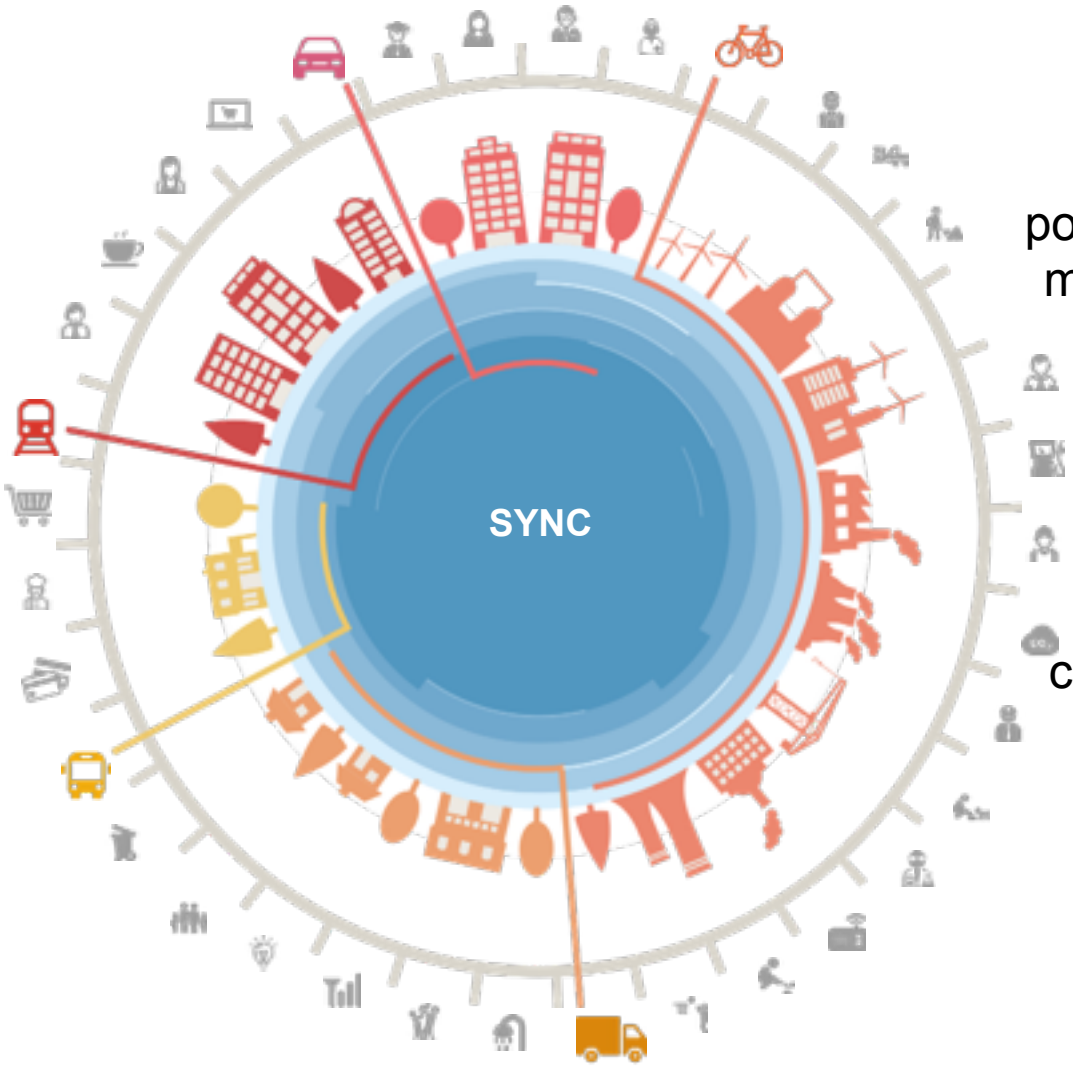
What is SYNC?



The utility of an interaction is a function of the costs and benefits of undertaking the journey and the costs and benefits of the opportunities at the other end

Interactions are not necessarily limited to physical human movements. They can equally be related to the costs and benefits of moving water, energy, waste and even information through utilities and communication networks

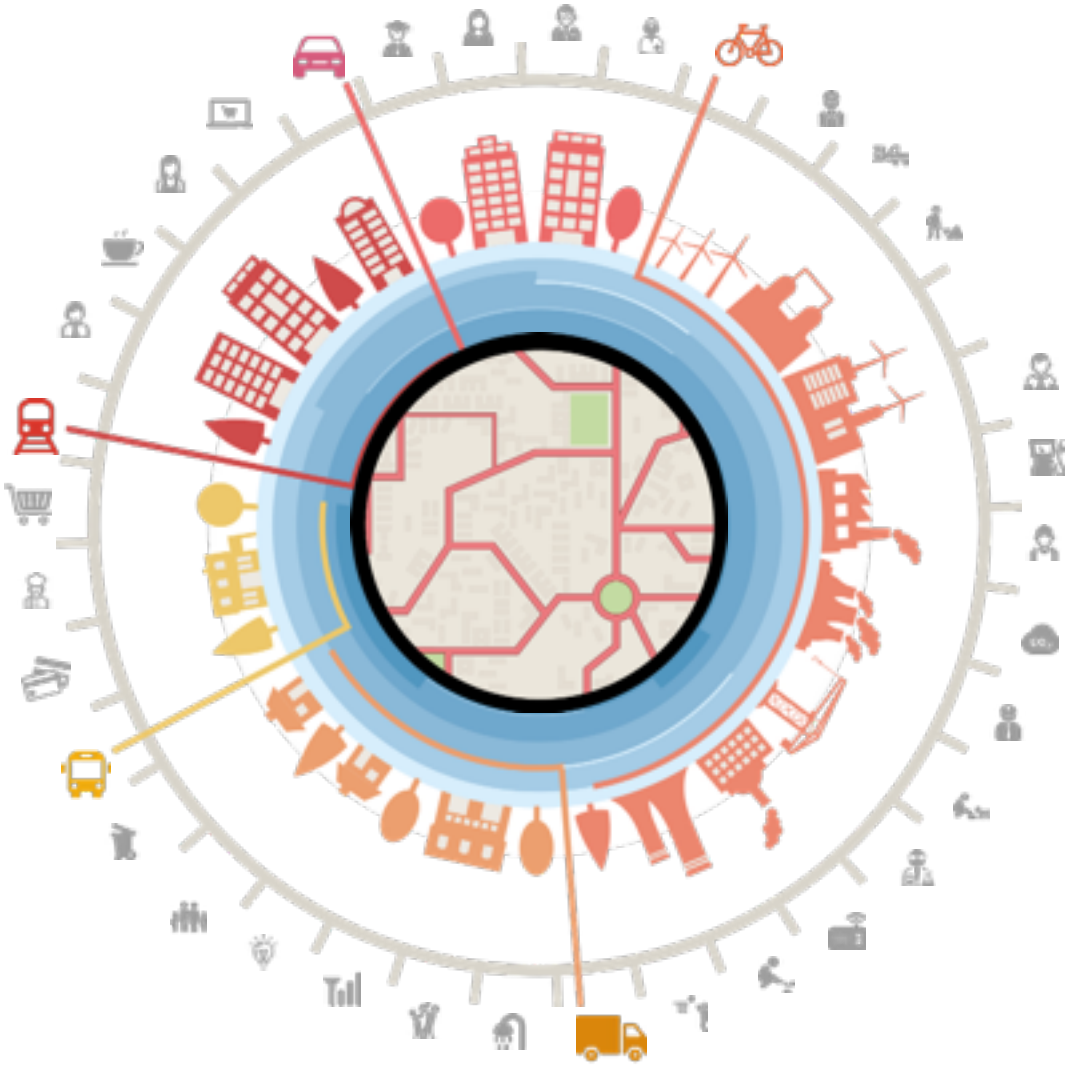
What is SYNC?



By understanding the relationships between activity, infrastructure, travel and interaction it is possible to generate a wide range of performance metrics of value in the strategic planning process

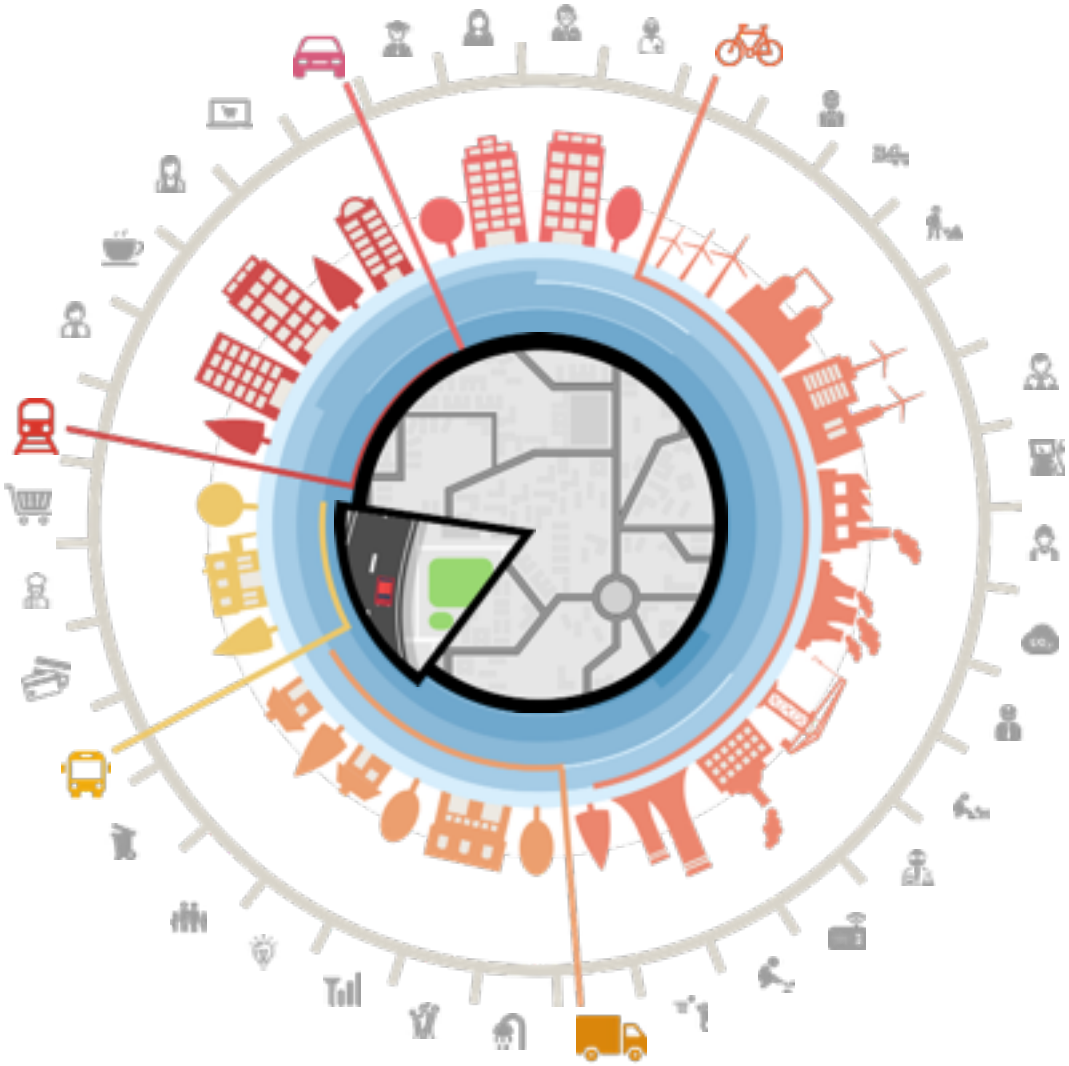
These include measures of travel demand, distribution, congestion, measures of location characteristics such as land use and value and measures of system resilience and efficiency, measures of location potential and pressure, consumption and production, growth and decline

What is SYNC?



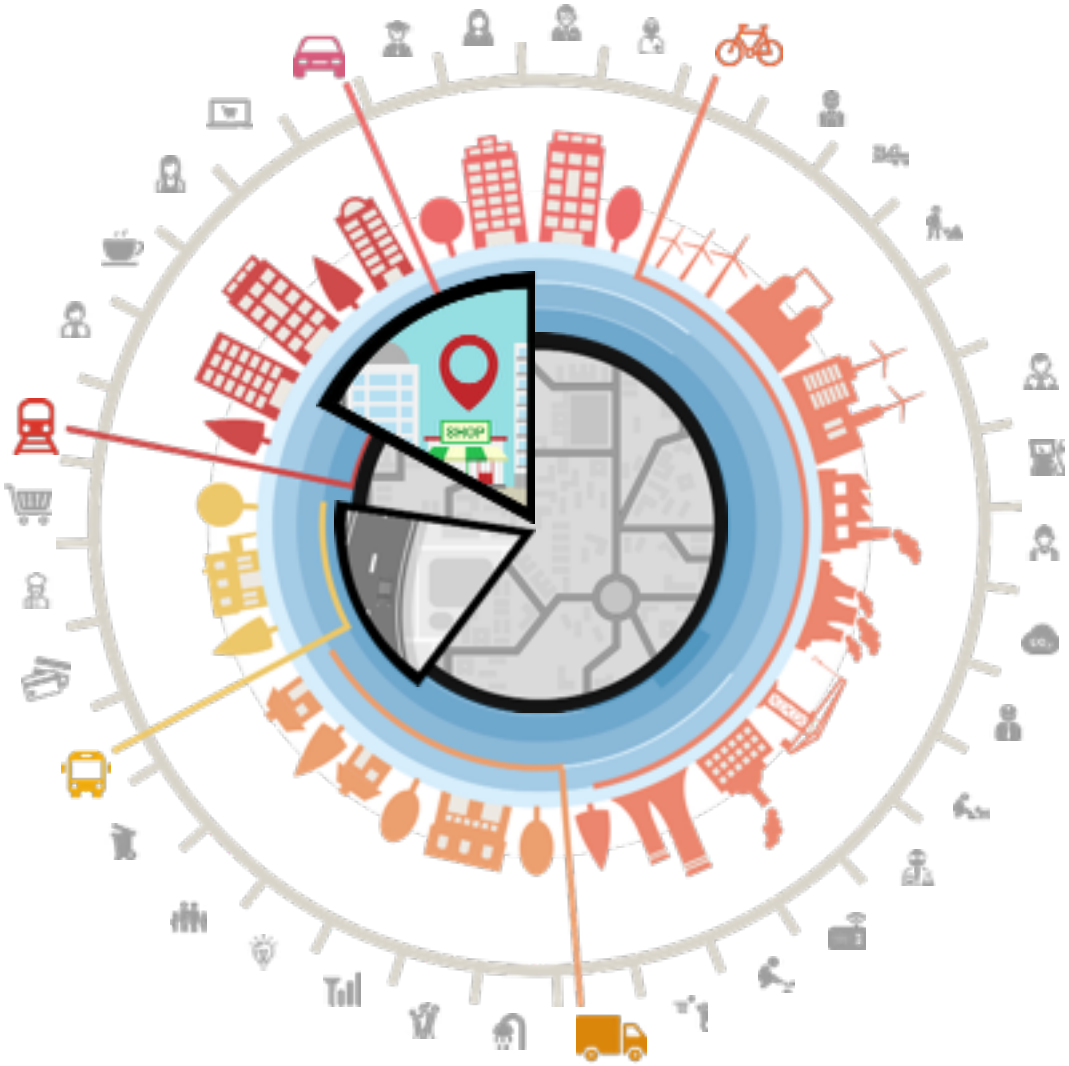
A distinctive feature of the SYNC platform is the connection made between interactions operating across geographic scales. Model applications embedded in SYNC can provide outputs at a city-regional level...

What is SYNC?



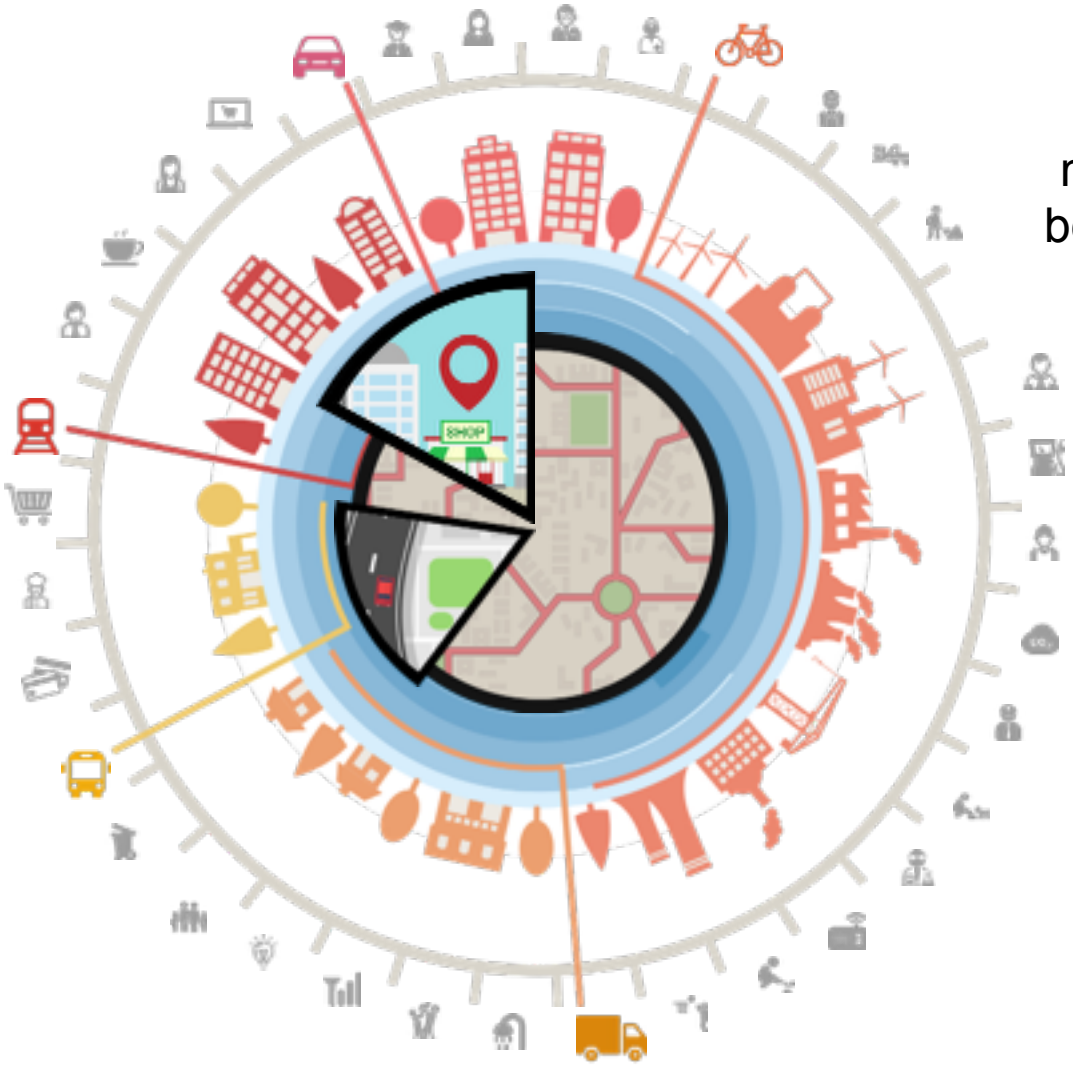
...at the neighbourhood level...

What is SYNC?



...and even at the building and plot level.

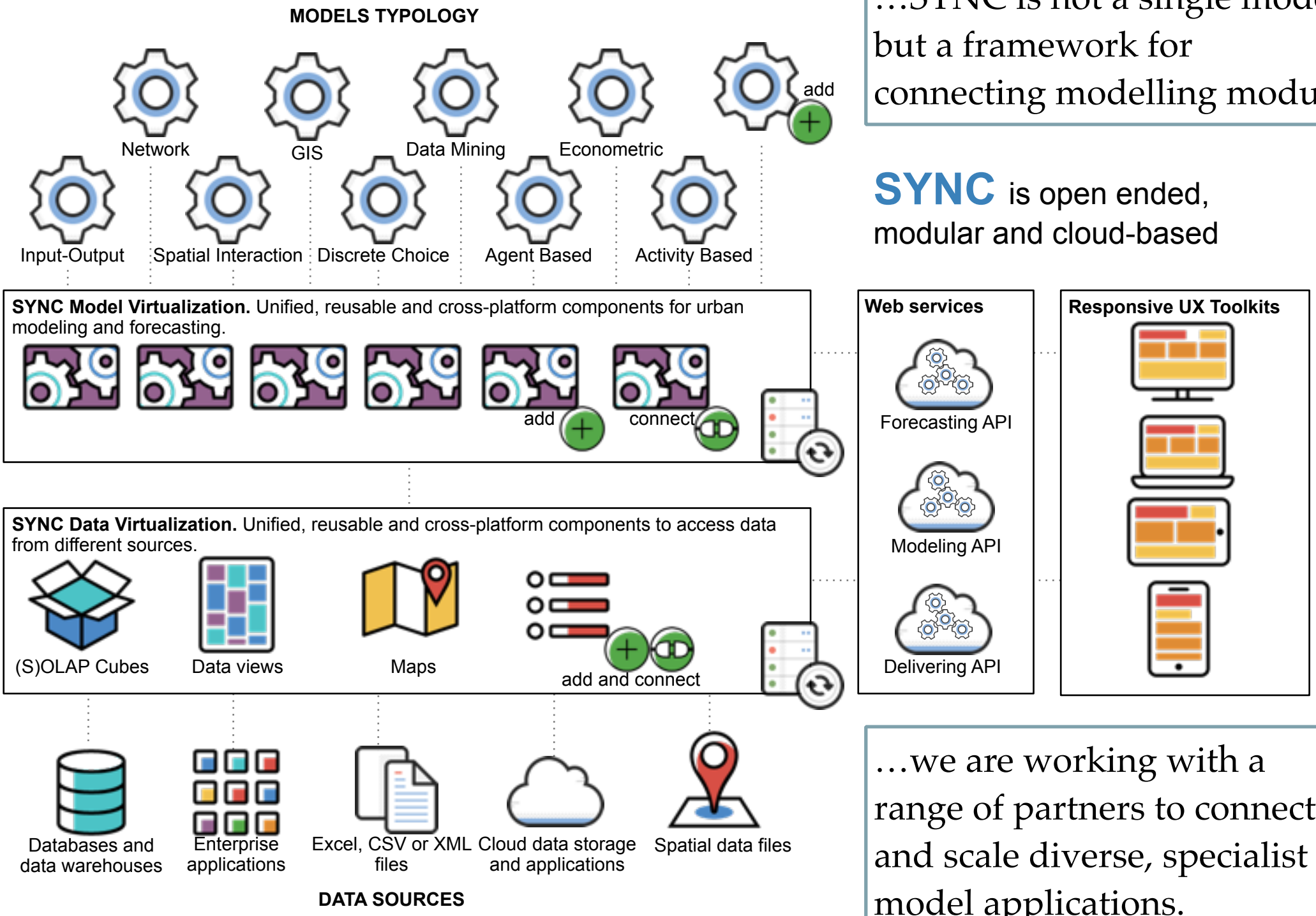
What is SYNC?



The interconnection between applications designed for different resolutions of evaluation means that the critical connections can be made between local or borough level decisions and city wide impacts (and vice versa).

What will SYNC look like?

What will SYNC look like?



…SYNC is not a single model but a framework for connecting modelling modules

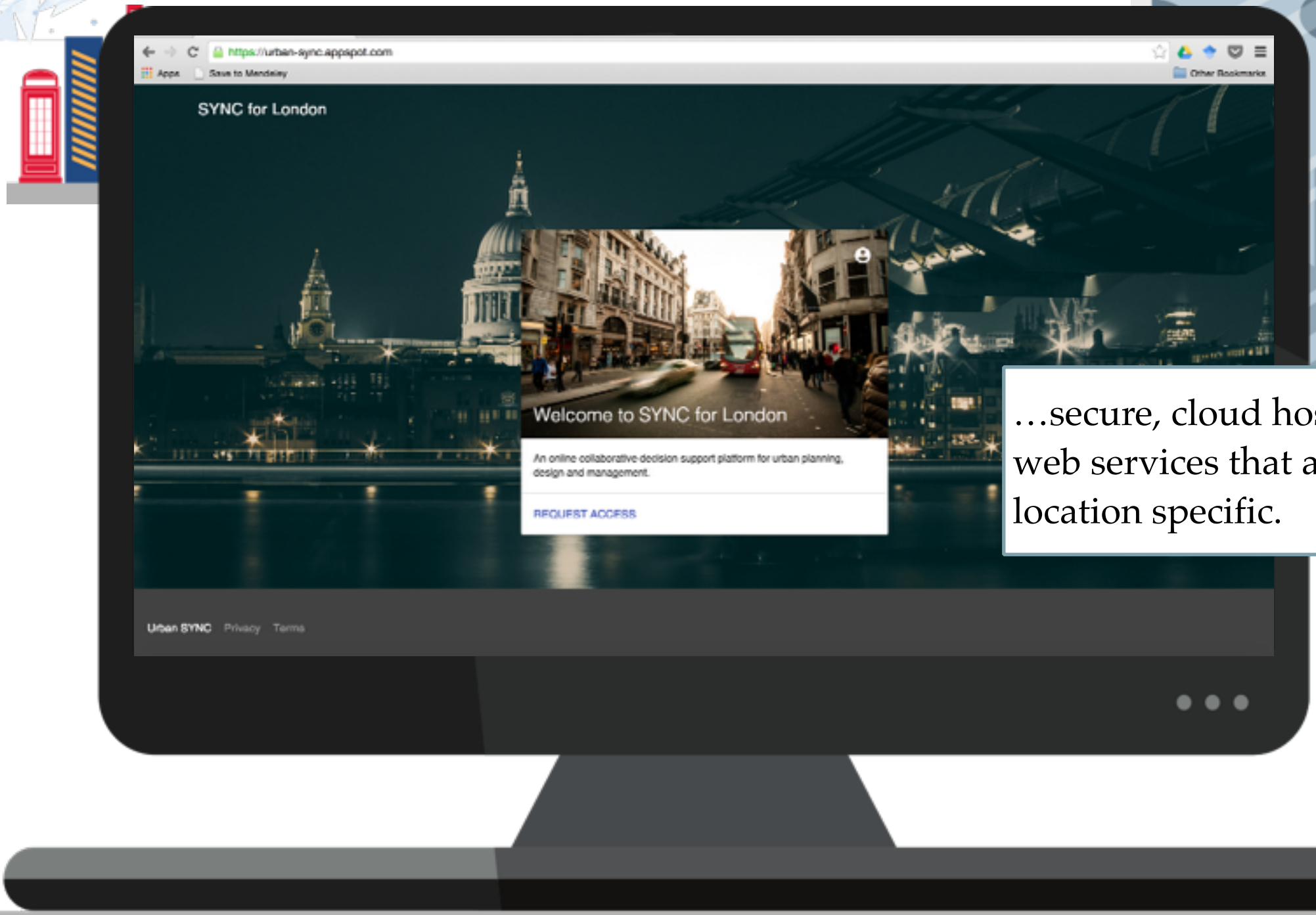
SYNC is open ended, modular and cloud-based

…we are working with a range of partners to connect and scale diverse, specialist model applications.

What will SYNC look like?

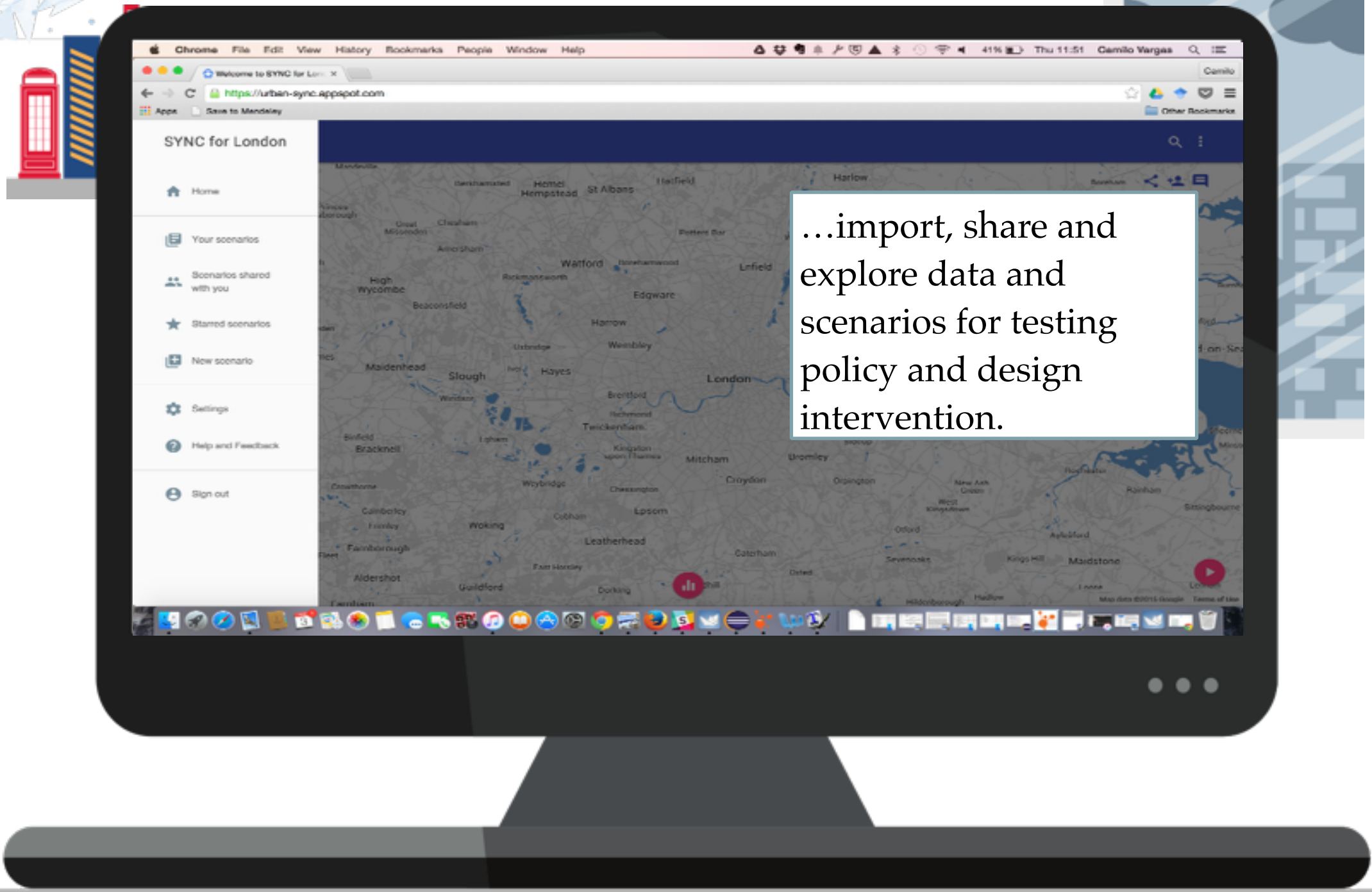
We aim to support different ways of interacting with SYNC resources from mobile access to hands on interactive displays.

What will SYNC look like?

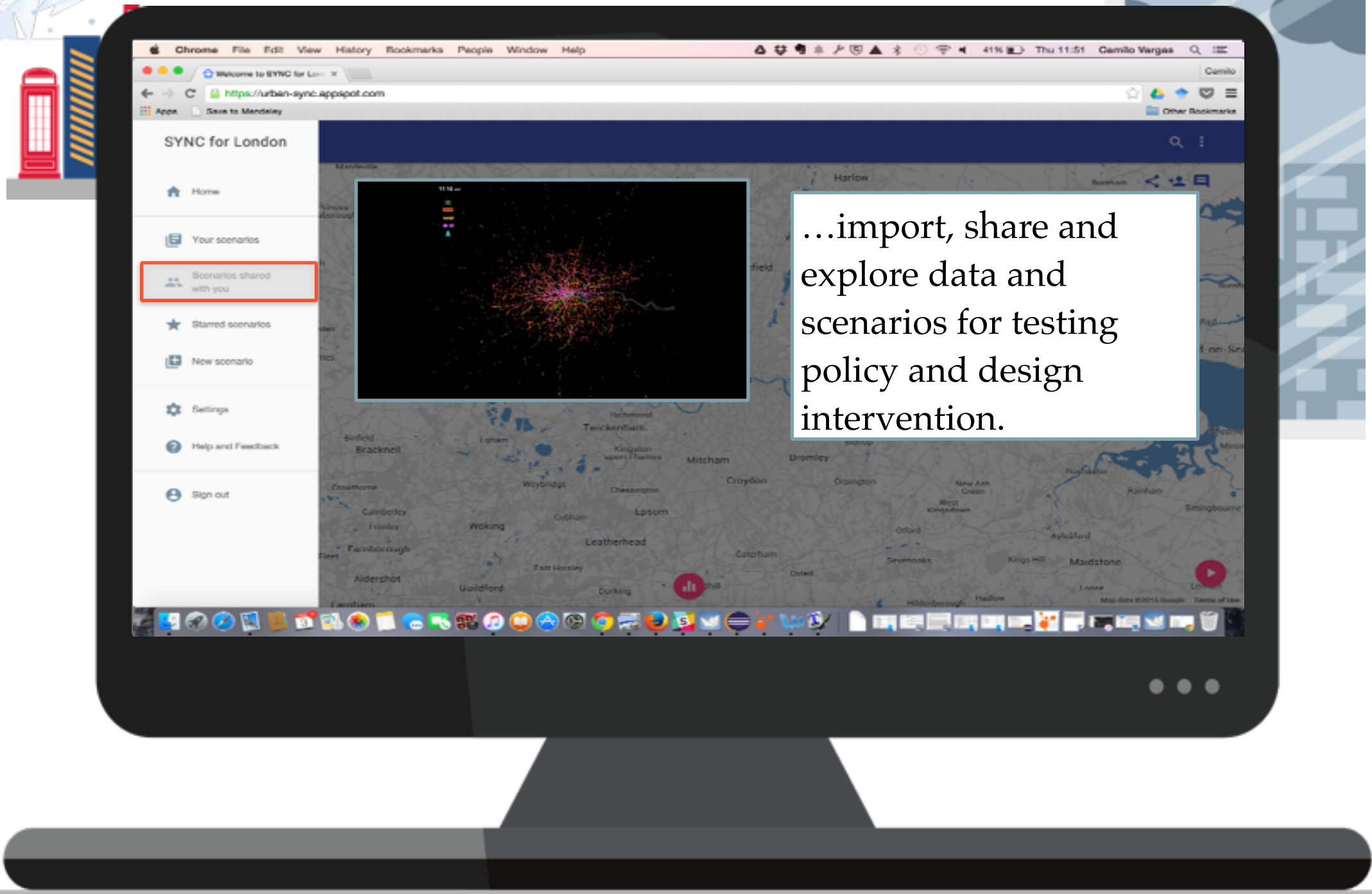


…secure, cloud hosted web services that are location specific.

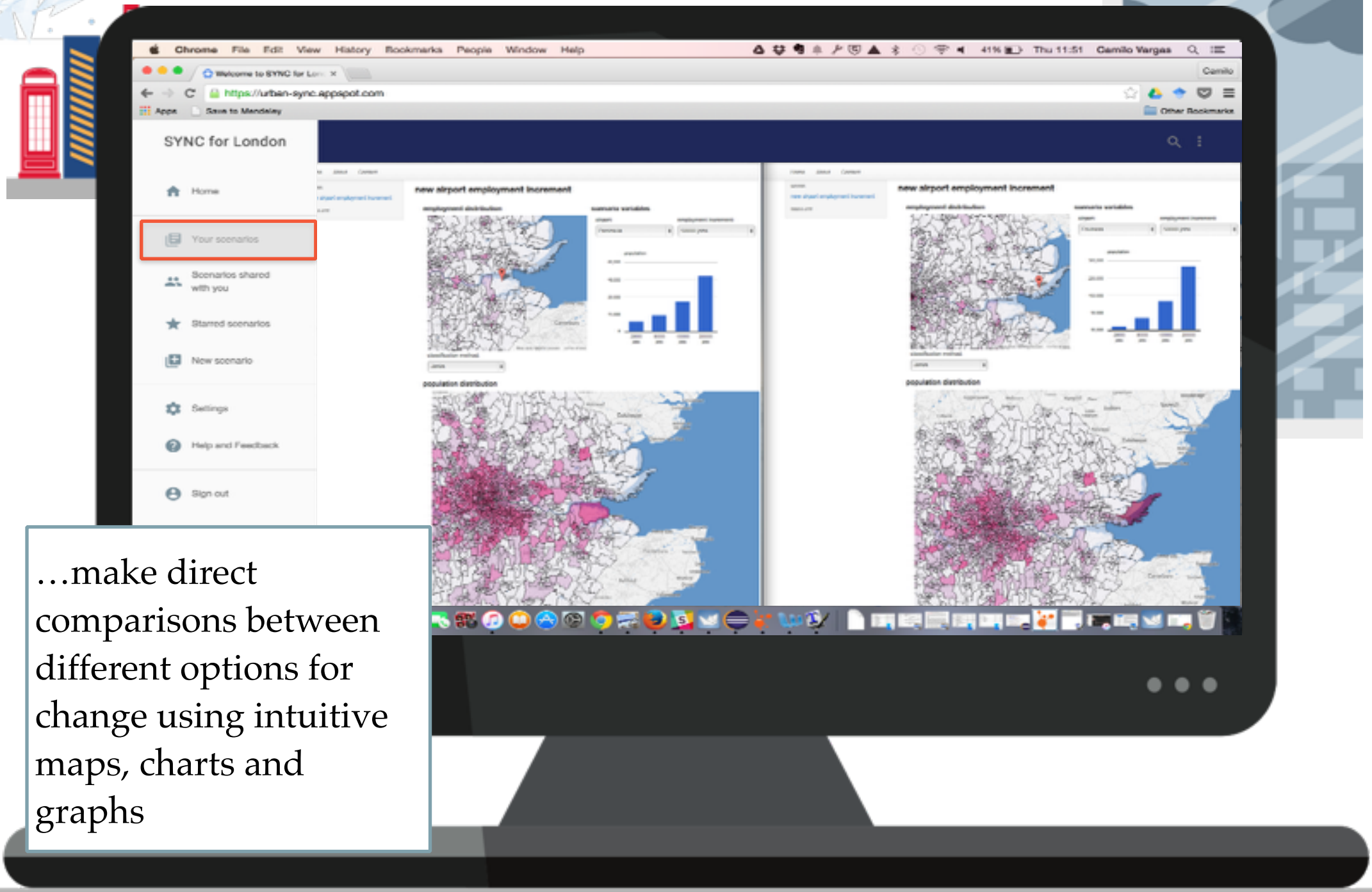
What will SYNC look like?



What will SYNC look like?

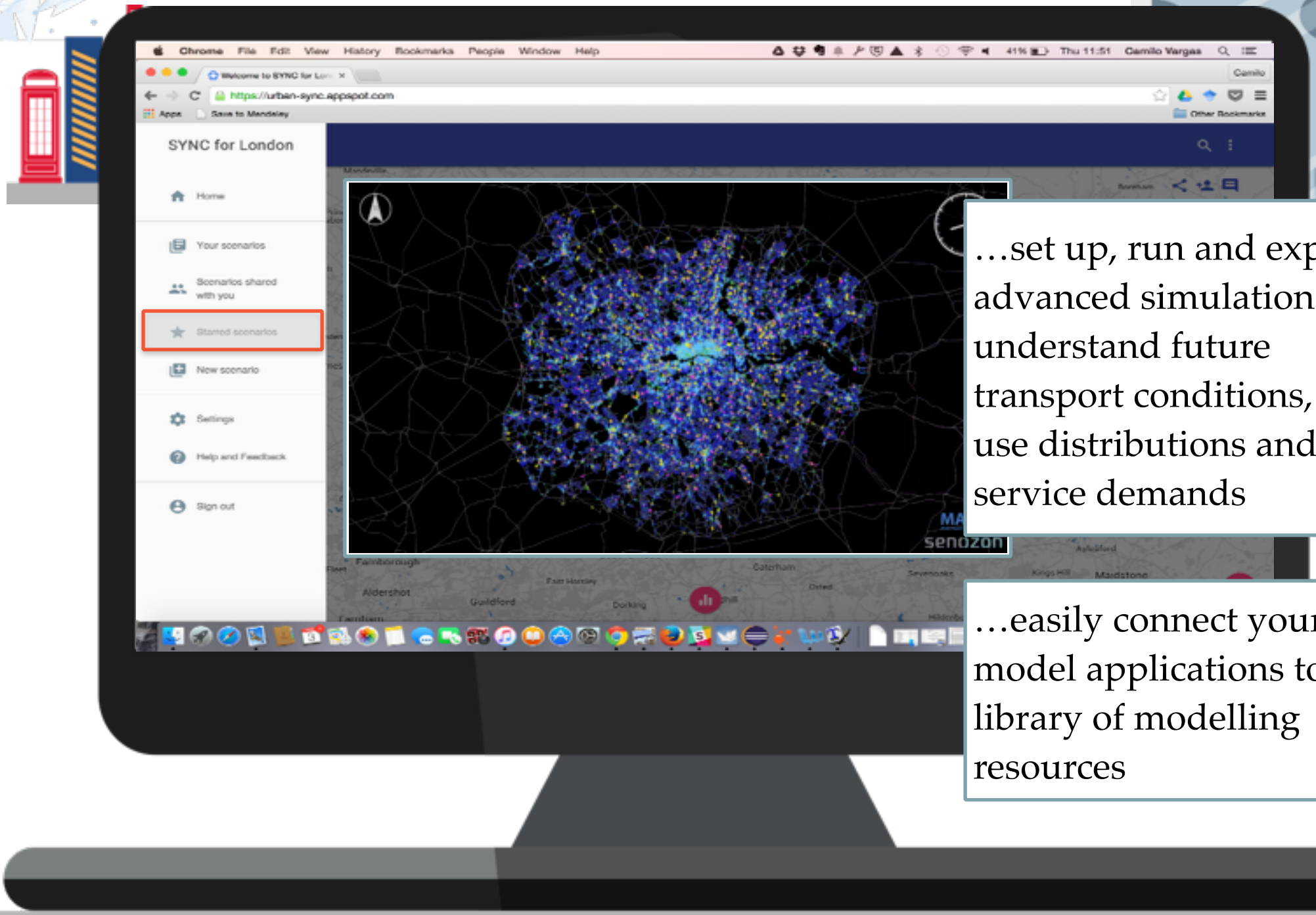


What will SYNC look like?



…make direct comparisons between different options for change using intuitive maps, charts and graphs

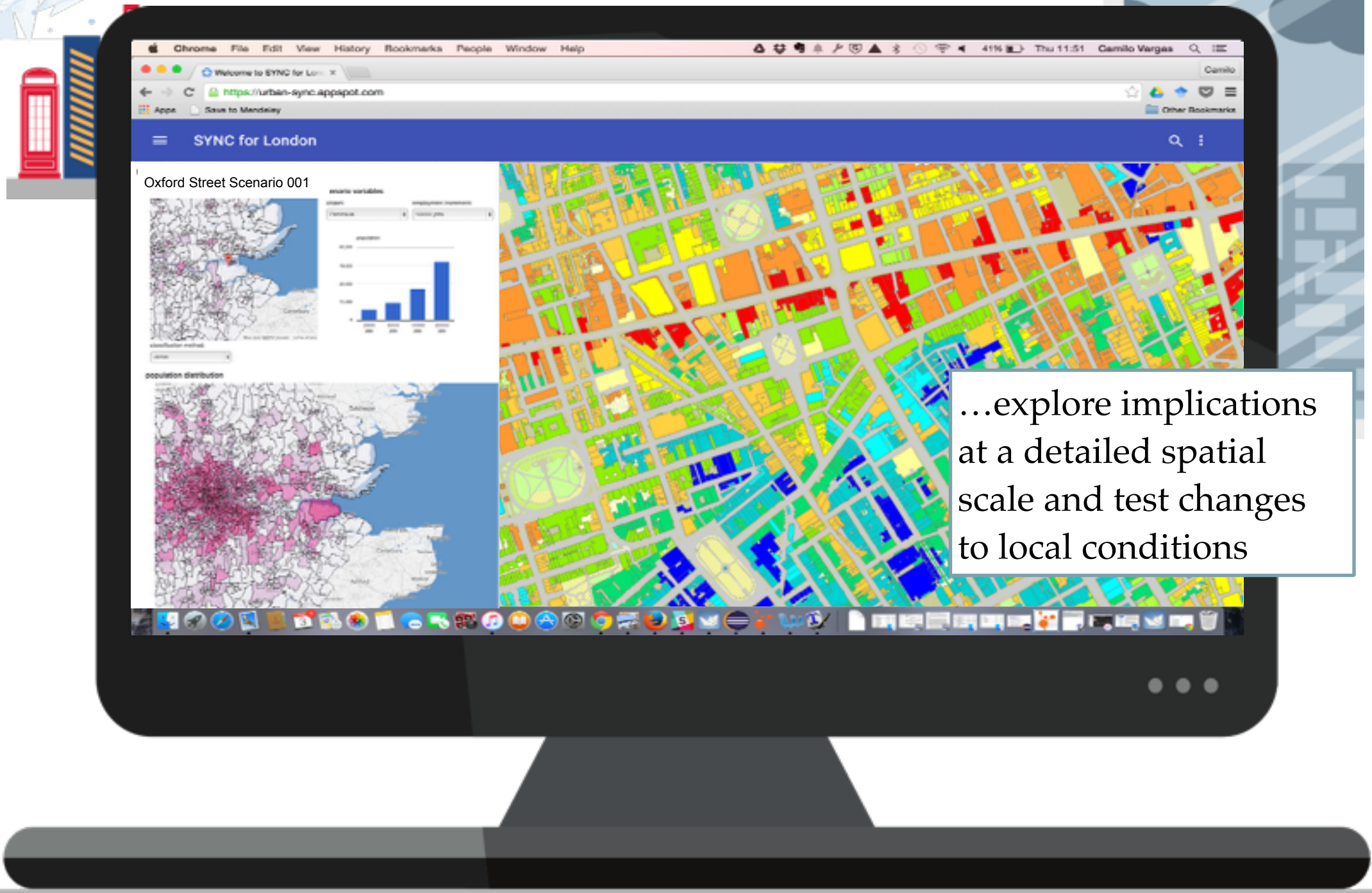
What will SYNC look like?



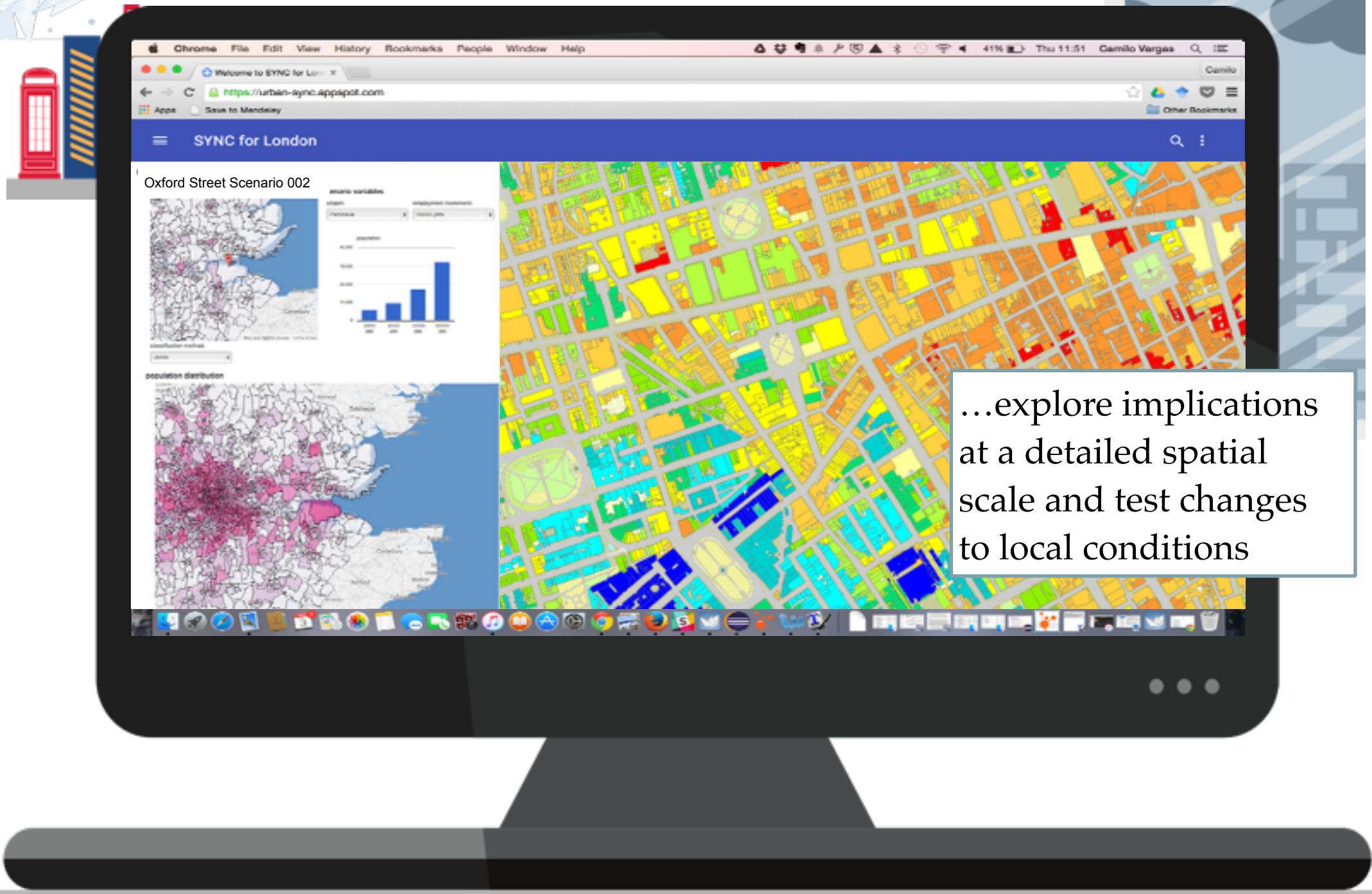
…set up, run and explore advanced simulations to understand future transport conditions, land use distributions and service demands

…easily connect your own model applications to this library of modelling resources

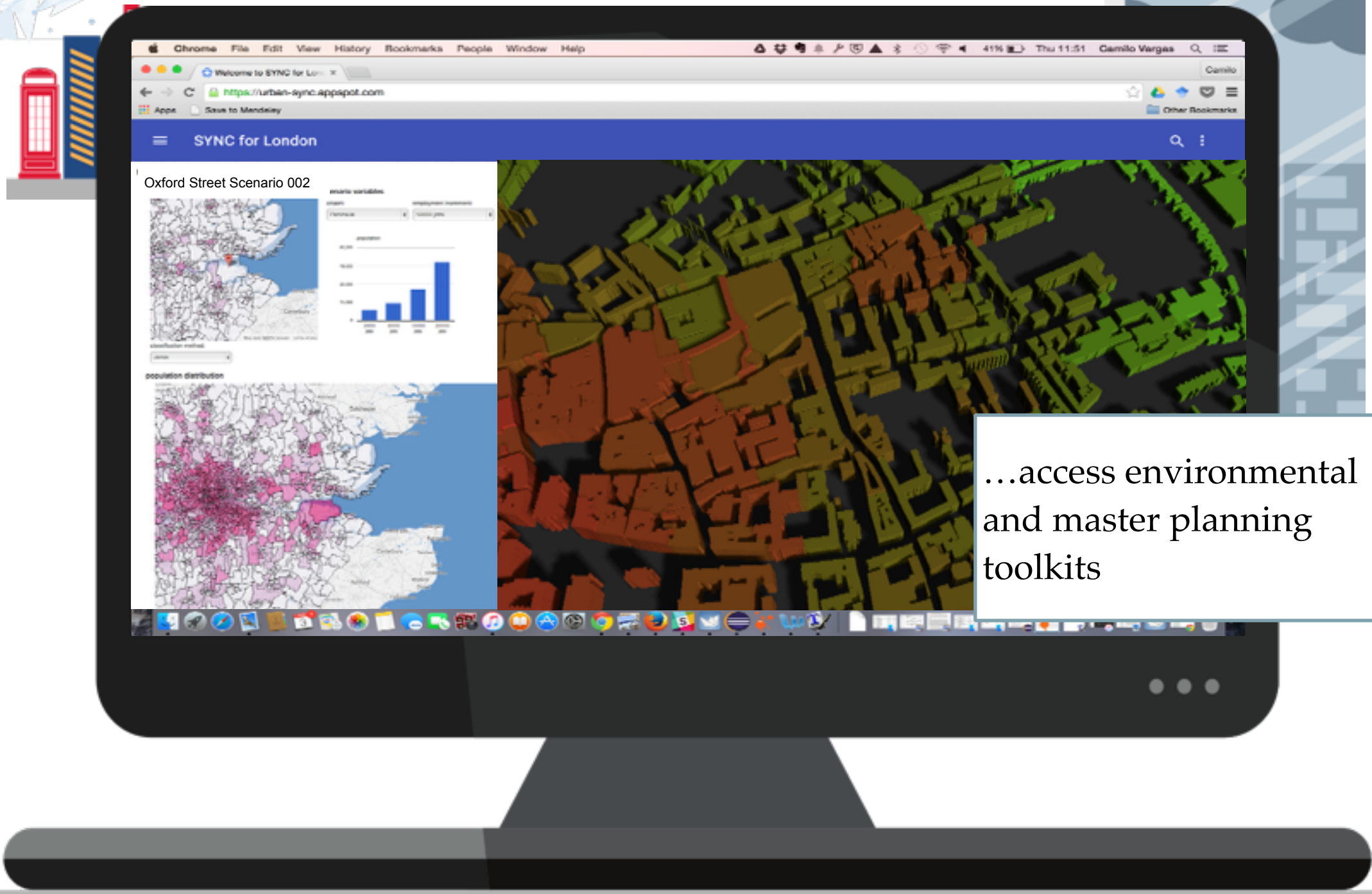
What will SYNC look like?



What will SYNC look like?



What will SYNC look like?



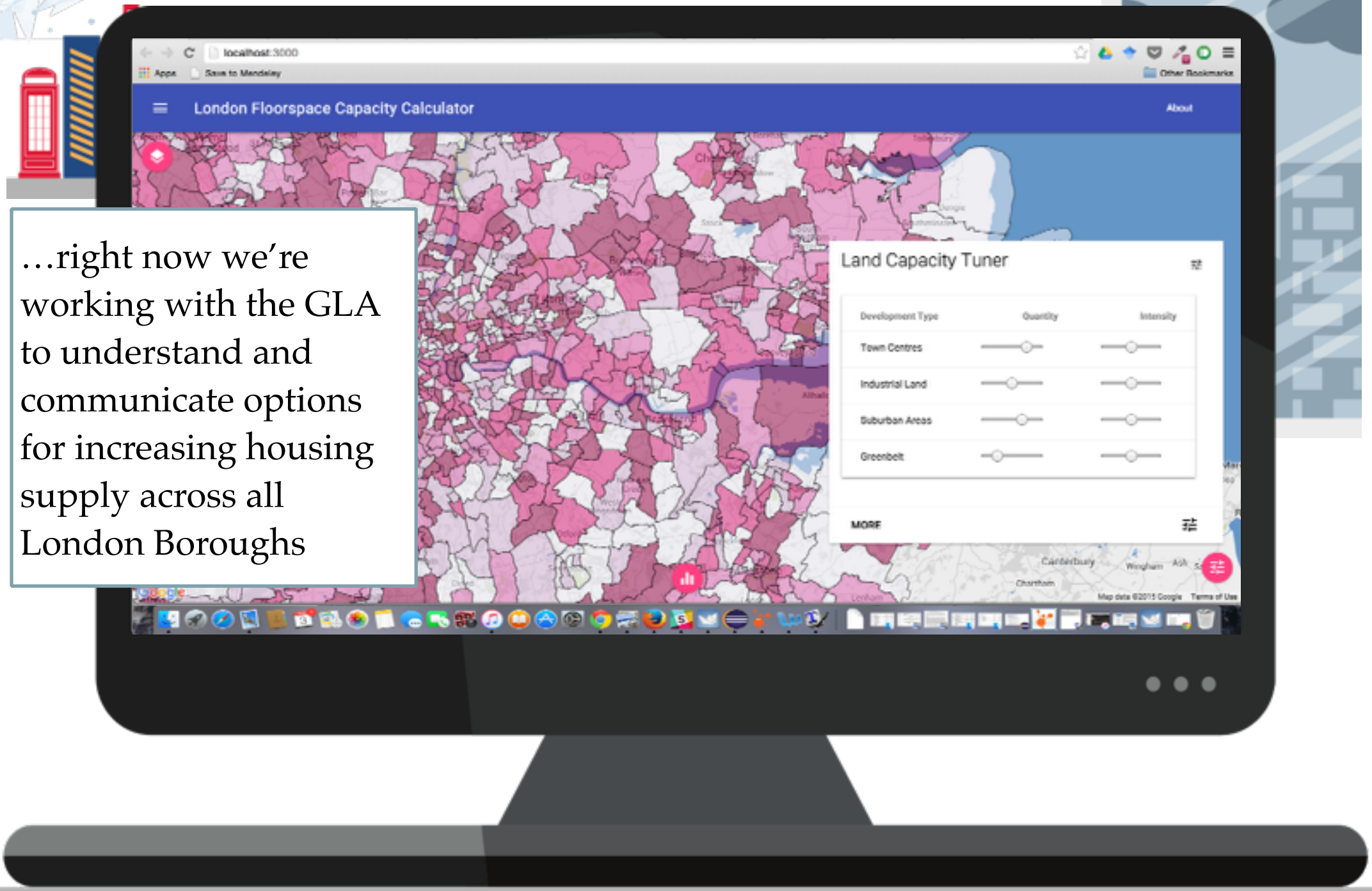
…access environmental and master planning toolkits

What are we doing now?

SYNC

What are we doing now?

…right now we're working with the GLA to understand and communicate options for increasing housing supply across all London Boroughs



we need your input!

…use cases
…common challenges
…specifications

We’re able to use our funding to support Borough level planning challenges. Let us know if you have a challenge that you think this resource might be able help with.

Thank you

pete.ferguson.10@ucl.ac.uk
pete@prospective.io

joan.serras@ucl.ac.uk
joan@prospective.io

London Borough Data Partnership Meeting #4

Data Science Institute, 7 October 2015

Attendance List

No.	First Name	Last Name	Organisation
1	Ahmed	Omer	Royal Borough of K&C
2	Alan	Lewis	Greater London Authority
3	Andrew	Collinge	Greater London Authority
4	Apollo	Gerolymbos	London Fire Brigade
5	C	Williams	Wandsworth
6	Daniel	Quirke	London Councils
7	David	Clifford	Islington
8	Denise	Bushay	Croydon
9	Fiona	O'Toole	Richmond
10	Indu	Srikumar	Redbridge
11	Jamie	Dickie	Bexley
12	Joan	Serras	UCL
13	Joshan	Meenowa	Cabinet Office
14	Kieran	O'Sullivan	Transport for London
15	Kim	Overy	Hillingdon
16	Lauren	Payne	Waltham Forest
17	Man Lim	Li	Waltham Forest
18	Marc	Marson	Harrow
19	Marnie	Caton	Islington
20	Mary-Ann	Domman	London Councils
21	Nerida	Devane	Greater London Authority
22	Owen	Kennedy	Hounslow
23	Paul	Hodgson	Greater London Authority
24	Rob	White	Waltham Forest
25	Ronan	Smyth	Lewisham

26	Simon	Pitman	Transport for London
27	Stuart	Carter	Southwark
28	Sudip	Trivedi	Camden
29	Victoria	Beard	Croydon
30	James	Leach	Southwark